

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2004, 00:26:49 ; Search time 4909 Seconds  
(without alignments)  
11045.454 Million cell updates/sec

Title: US-10-603-141-1  
Perfect score: 1251  
Sequence: 1 gaatcggcagcaggcaga.....gtatcaggagacgtgcac 1251

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.ov.\*
- 22: em.pat.\*
- 23: em.ph.\*
- 24: em.pl.\*
- 25: em.ro.\*
- 26: em.sts.\*
- 27: em.un.\*
- 28: em.vi.\*
- 29: em.htg.hum.\*
- 30: em.htg.inv.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sv.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1251	100.0	1251	10	AF301018 Mus muscu
2	1229.8	98.3	1823	10	AF305709 Mus muscu
3	1053	84.2	1053	6	E28986 Seven-pass
4	1053	84.2	1053	6	E28987 Seven-pass
5	1009.4	80.7	231959	2	AC126907 Rattus no
6	1009.4	80.7	254420	2	AC107545 Rattus no
7	709.6	56.7	2804	9	HSU73531 Human G pro
8	706.4	56.5	1923	9	BC033584 Homo sapi
9	706.4	56.5	2238	6	AR049107 Sequence
10	705.4	56.4	1953	6	AX549062 Sequence
11	705.4	56.4	1953	6	HSY13248 Homo sapien
12	704.8	56.3	1918	9	HSU73529 Human G pro
13	704.8	56.3	209772	9	AC099782 Homo sapi
14	704.8	56.3	219553	2	HSX312887 Homo sapi
15	696	55.6	1032	9	AF124380 Macaca mu
16	696	55.6	1032	9	AF237559 Cercopitu
17	694.4	55.5	1072	9	AF291671 Macaca fa
18	689.4	55.1	1029	6	AR158353 Sequence
19	689.4	55.1	1029	6	AF007858 Macaca ne
20	685	54.8	1037	6	AR158352 Sequence
21	685	54.8	1037	9	AF007859 Cercopith
22	679.2	54.3	1029	6	AR178364 Sequence
23	679.2	54.3	1029	9	AF007545 Homo sapi
24	679.2	54.3	1029	9	AF029759 Homo sapi
25	679.2	54.3	1029	9	AF084229 Pan trogl
26	679.2	54.3	1029	6	AY322543 Homo sapi
27	679.2	54.3	1029	6	AX244235 Sequence
28	676.8	54.1	1026	6	AR151483 Sequence
29	676.8	54.1	1026	6	AX015624 Sequence
30	648	51.8	1026	6	AR158351 Sequence
31	304.8	24.4	449	9	AF354630 Gorilla g
32	300.2	24.0	449	9	AF354631 Pongo pyg
33	184.4	14.7	215176	2	BX546447 Danio rer
34	183.4	14.7	1074	6	AR308572 Sequence
35	183.4	14.7	1074	9	AY242127 Homo sapi
36	183.4	14.7	1186	9	HSX132337 Homo sapi
37	183.4	14.7	2462	6	AX548955 Sequence
38	183.4	14.7	2462	9	AF145440 Homo sapi
39	183.4	14.7	2544	9	AF145439 Homo sapi
40	183.4	14.7	2577	6	BD271525 Anti-GPR-
41	183.4	14.7	2577	6	AR270532 Sequence
42	183.4	14.7	2577	6	AR366325 Sequence
43	183.4	14.7	2577	6	AX454125 Sequence
44	183.4	14.7	2577	9	HSU45982 Human G pro
45	183.4	14.7	176968	9	AC005669 Homo sapi

ALIGNMENTS

RESULT 1  
AF301018  
LOCUS AF301018 1251 bp mRNA linear ROD 25-NOV-2000  
DEFINITION Mus musculus strain C57BL/6 chemokine receptor CXCR6 mRNA, complete cds.  
ACCESSION AF301018  
VERSION AF301018.1 GI:11342661  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1251)  
AUTHORS Matloubian,M., David,A., Engel,S., Ryan,J.E. and Cyster,J.G.  
TITLE A transmembrane CXC chemokine is a ligand for HIV-coreceptor Bonzo





```
QY 78 ATGATGATGGGCAATCAAGAGTACGCTCTGTACGATGGGCACTACGAGGAGATTTCTGG 137
Db 1 ATGATGATGGGCAATCAAGAGTACGCTCTGTACGATGGGCACTACGAGGAGATTTCTGG 60
QY 138 CTCCTTCAACAATTCAGTGATAACAGCAGCAGAGAAACAAACGCTTCTTAAAGTTCAAGGAG 197
Db 61 CTCCTTCAACAATTCAGTGATAACAGCAGCAGAGAAACAAACGCTTCTTAAAGTTCAAGGAG 120
QY 198 GTCTTTTGGCCCTGTGTACTCCTGGTAGTGTGTCTTGGAGCTGTAGAAACTCCCTG 257
Db 121 GTCTTTTGGCCCTGTGTACTCCTGGTAGTGTGTCTTGGAGCTGTAGAAACTCCCTG 180
QY 258 GTTCTGATTATATACATTTCTACCAAGCTGAGGACTCTGACAGATGTGTTCTGCTG 317
Db 181 GTTCTGATTATATACATTTCTACCAAGCTGAGGACTCTGACAGATGTGTTCTGCTG 240
QY 318 AACTTGGCCCTGGCTGACCTGGTGTGTGTCTGTACTCTGCCCCTTTTGGGCTATGCAGGC 377
Db 241 AACTTGGCCCTGGCTGACCTGGTGTGTGTCTGTACTCTGCCCCTTTTGGGCTATGCAGGC 300
QY 378 ACCTATGAGTGGGTCTTTGGCACAGTCAATGTGCAAACTCTTCGAGGCATGTATACAATG 437
Db 301 ACCTATGAGTGGGTCTTTGGCACAGTCAATGTGCAAACTCTTCGAGGCATGTATACAATG 360
QY 438 AACTTCTACGTGTCCATGCTCACTCTCACCTGCACTACAGTGGATCGTTTCATTGTAGTG 497
Db 361 AACTTCTACGTGTCCATGCTCACTCTCACCTGCACTACAGTGGATCGTTTCATTGTAGTG 420
QY 498 GTCCAGGCTACCAAGGCTTCAACCGCAGGCTAAGTGAAGATCTGGGCGCCAAAGTCATT 557
Db 421 GTCCAGGCTACCAAGGCTTCAACCGCAGGCTAAGTGAAGATCTGGGCGCCAAAGTCATT 480
QY 558 TGGTGTCTCATTTGGGTGGTCTCCCTCTGGTCTCTTGGTCTTGGTCCACAGATCACTATGCCAT 617
Db 481 TGGTGTCTCATTTGGGTGGTCTCCCTCTGGTCTCTTGGTCTTGGTCCACAGATCACTATGCCAT 540
QY 618 GTTCAAGATATTGAAGCTTATCTGTCAAGTACACAGTGGAGGATATCCACTATGGTT 677
Db 541 GTTCAAGATATTGAAGCTTATCTGTCAAGTACACAGTGGAGGATATCCACTATGGTT 600
QY 678 CTTGTATACAGATGACTCTGGGTCTTCTTCTGCAATGTCTCACTATGATCTTGTGCTAC 737
Db 601 CTTGTATACAGATGACTCTGGGTCTTCTTCTGCAATGTCTCACTATGATCTTGTGCTAC 660
QY 738 TCAGGCATTATCAAGACCTTGCTTCATGCTCGAACTTCCAGAACGACAAATCTCTAAAG 797
Db 661 TCAGGCATTATCAAGACCTTGCTTCATGCTCGAACTTCCAGAACGACAAATCTCTAAAG 720
QY 798 ATCATCTTCCCTTGTAGTGGTGTCTTCTGCTGACCCAGACACCCCTTCAACCTTGGCATG 857
Db 721 ATCATCTTCCCTTGTAGTGGTGTCTTCTGCTGACCCAGACACCCCTTCAACCTTGGCATG 780
QY 858 TTAATCCAAAGTACAGCTGGGCTGCTTACCAATACCAAGCTTAAAGTATGCCATCGTA 917
Db 781 TTAATCCAAAGTACAGCTGGGCTGCTTACCAATACCAAGCTTAAAGTATGCCATCGTA 840
QY 918 GTGACAGAGGCTATAGCATATTTCCGGCTTGGCTTAAACCTGTACTTTATGCTTTGTT 977
Db 841 GTGACAGAGGCTATAGCATATTTCCGGCTTGGCTTAAACCTGTACTTTATGCTTTGTT 900
QY 978 GGCTTAAAGTCCGAGAGAGCTCTGAAACTTATGAAGTATCGGCTGCTCTCAC 1037
Db 901 GGCTTAAAGTCCGAGAGAGCTCTGAAACTTATGAAGTATCGGCTGCTCTCAC 960
QY 1038 CTGGAGTCTCAAGTCAATGAAAGTCTTCTGAGGACAGTTCGAAGATTTGTTCTGCTCC 1097
Db 961 CTGGAGTCTCAAGTCAATGAAAGTCTTCTGAGGACAGTTCGAAGATTTGTTCTGCTCC 1020
QY 1098 CACAATGTAGAGACACCAAGTATGTTCAATTG 1130
Db 1021 CACAATGTAGAGACACCAAGTATGTTCAATTG 1053
```

```
RESULT 4
E28987/c
LOCUS
DEFINITION
E28987
Seven-pass transmembrane receptor protein ET60.
ACCESSION
E28987.1
GI:13026578
VERSION
JP 1999155573-A/2.
KEYWORDS
unidentified
SOURCE
unidentified
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 1053)
AUTHORS
Hiroshi, I., Naoko, I. and Takehiro, K.
TITLE
Seven-pass transmembrane receptor protein ET60
JOURNAL
Patent; JP 1999155573-A 2 15-JUN-1999;
ASAHI CHEM IND CO LTD
OS
Unidentified
PN
JP 1999155573-A/2
PD
15-JUN-1999
PF
27-NOV-1997 JP 1997325823
PR
HIROSHI ISHIMARU, NAOKO IEMURA, TAKEHIRO KOSHIO PC
C12N15/09, C07K14/705, C07K16/28, C12N5/10, C12P21/02, G01N33/566, PC
C12N15/00,
PC
C12N5/00
CC
Strandedness: Single;
CC
Topology: Linear;
FH
Key Location/Qualifiers
FT
source 1..1053
/organism="Unidentified".
FEATURES
source
Location/Qualifiers
1..1053
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 84.2%; Score 1053; DB 6; Length 1053;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 78 ATGATGATGGGCAATCAAGAGTACGCTCTGTACGATGGGCACTACGAGGAGATTTCTGG 137
Db 1053 ATGATGATGGGCAATCAAGAGTACGCTCTGTACGATGGGCACTACGAGGAGATTTCTGG 994
QY 138 CTCCTTCAACAATTCAGTGATAACAGCAGCAGAGAAACAAACGCTTCTTAAAGTTCAAGGAG 197
Db 993 CTCCTTCAACAATTCAGTGATAACAGCAGCAGAGAAACAAACGCTTCTTAAAGTTCAAGGAG 934
QY 198 GTCTTTTGGCCCTGTGTACTCTGCTGTAGTGTGTCTTGGAGCTGTAGGAACTCCCTG 257
Db 933 GTCTTTTGGCCCTGTGTACTCTGCTGTAGTGTGTCTTGGAGCTGTAGGAACTCCCTG 874
QY 258 GTTCTGATTATATACATTTCTACCAAGCTGAGGACTCTGACAGATGTGTTTCTGCTG 317
Db 873 GTTCTGATTATATACATTTCTACCAAGCTGAGGACTCTGACAGATGTGTTTCTGCTG 814
QY 318 AACTTGGCCCTGGCTGACCTGGTGTGTGTGTACTCTGCCCCTTTTGGGCTATGCAGGC 377
Db 813 AACTTGGCCCTGGCTGACCTGGTGTGTGTGTACTCTGCCCCTTTTGGGCTATGCAGGC 754
QY 378 ACCTATGAGTGGTCTTTTGGCACAGTCAATGTGCAAACTCTTCGAGGCACTATACAATG 437
Db 753 ACCTATGAGTGGTCTTTTGGCACAGTCAATGTGCAAACTCTTCGAGGCACTATACAATG 694
QY 438 AACTTCTACGTGTCCATGCTCACTCTCACCTGCACTACAGTGGATCGTTTCATTGTAGTG 497
Db 693 AACTTCTACGTGTCCATGCTCACTCTCACCTGCACTACAGTGGATCGTTTCATTGTAGTG 634
QY 498 GTCCAGGCTACCAAGGCTTCAACCGCAGGCTCAAGTGAAGATCTGGGCGCCAAAGTCATT 557
Db 633 GTCCAGGCTACCAAGGCTTCAACCGCAGGCTCAAGTGAAGATCTGGGCGCCAAAGTCATT 574
```

```

QY 558 TGCTTGCTCATTTGGTGGTCTCCCTGTTGGTTCCTTTGCGCACAGATCATCTATGGCCAT 617
Db 573 TGCTTGCTCATTTGGTGGTCTCCCTGTTGGTTCCTTTGCGCACAGATCATCTATGGCCAT 514
QY 618 GTTCAAGATATTGACAAGCTTATCTGTCAGTACACACAGTGGAGAGATATCCACTATGGTT 677
Db 513 GTTCAAGATATTGACAAGCTTATCTGTCAGTACACACAGTGGAGAGATATCCACTATGGTT 454
QY 678 CTTGTTATACAGATGACTCTGGGGTTCCTTCTGCGCATGTCCTCATATGATCTGTCGTAC 737
Db 453 CTTGTTATACAGATGACTCTGGGGTTCCTTCTGCGCATGTCCTCATATGATCTGTCGTAC 394
QY 738 TCAGGCATTATCAAGACCTTGCTTCATGCTCGCAAACTTCCAGAAGCAAAATCTCTAAAG 797
Db 393 TCAGGCATTATCAAGACCTTGCTTCATGCTCGCAAACTTCCAGAAGCAAAATCTCTAAAG 334
QY 798 ATCATCTTCCTTTAGTGGCTGTGTTCTGCTGTGACCCAGACACACCTTCAACCTTGGCATG 857
Db 333 ATCATCTTCCTTTAGTGGCTGTGTTCTGCTGTGACCCAGACACACCTTCAACCTTGGCATG 274
QY 858 TTAATCCAAAGTACAAAGCTGGAGTACTATACCATACCCAGTCTTAAAGTATGCCATGTA 917
Db 273 TTAATCCAAAGTACAAAGCTGGAGTACTATACCATACCCAGTCTTAAAGTATGCCATGTA 214
QY 918 GTGACAGAGGCTATAGCATACTTTTGGGCTTGCCCTTAAACCTGACTTTATGCGCTTTGTT 977
Db 213 GTGACAGAGGCTATAGCATACTTTTGGGCTTGCCCTTAAACCTGACTTTATGCGCTTTGTT 154
QY 978 GCGTTAAAGTTCGGAAGAGCTGCTGGAACATTAAGAGATATGCGCTGCTCTCCAC 1037
Db 153 GCGTTAAAGTTCGGAAGAGCTGCTGGAACATTAAGAGATATGCGCTGCTCTCCAC 94
QY 1038 CTGGGAGTCTCAAGTCAATGGAAGCTTCTGAGGACAGTTCGAAGACTTGTTCGCTCC 1097
Db 93 CTGGGAGTCTCAAGTCAATGGAAGCTTCTGAGGACAGTTCGAAGACTTGTTCGCTCC 34
QY 1098 CACAATGTAGACACACAGTATGTTCCAAATG 1130
Db 33 CACAATGTAGACACACAGTATGTTCCAAATG 1

RESULT 5
AC126907
LOCUS
DEFINITION
Rattus norvegicus clone CH230-150D22, WORKING DRAFT SEQUENCES, 6
unordered pieces.
ACCESSION
AC126907
VERSION
AC126907.3 GI:25007774
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus
REFERENCE
1. (bases 1 to 231959)
Muzny,D., Marie, Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Aman,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,N., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Chen,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Georgievski,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,

```

Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,  
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,  
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
 Kowis, C., Kratt, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorensuhewa, L., Loulseghe, H., Lozado, R.J., Lu, X., Ma, J.,  
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martine, E.,  
 Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,  
 Milosavljevic, A., Miner, G., Minja, S., Munidasa, M., Murphy, M., Nair, L.,  
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Nguyen, S.,  
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
 Nwakoelameh, O., Okwuonu, G., Olarnungagoon, A., Pal, S., Parks, K.,  
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
 Plopper, P., Poindexter, A., Popovic, D., Primus, E., Pu, L., L.,  
 Puzos, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,  
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
 Rives, C., Rodkey, T., Rojas, A., Rose, R., Rose, R., Ruiz, S.J.,  
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,  
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
 Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
 Valas, R., Vera, V., Villabana, D., Waldron, L., Walker, B., Wang, J.,  
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
 Williams, G., Willson, R., Wleczky, R., Woodson, H., Worley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von  
 Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
 Weinstein, G. and Gibbs, R.A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 231959)  
 Worley, K.C.  
 Direct Submission  
 Submitted (10-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 231959)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (15-NOV-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Nov 15, 2002 this sequence version replaced gi:23195284.  
 The sequence in this assembly is a combination of SAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: GZSX  
 Center clone name: CH230-150D22  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 219672 bases at least Q40  
 Consensus quality: 221104 bases at least Q30  
 Consensus quality: 222280 bases at least Q20  
 Estimated insert size: 218775; sum-of-contigs estimation  
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

```

*****
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 82701: contig of 82701 bp in length
* 82702 82801: gap of unknown length
* 82802 82802: contig of 129261 bp in length
* 212063 212062: contig of unknown length
* 212163 212162: gap of unknown length
* 215821 215820: contig of 3658 bp in length
* 215821 215820: gap of unknown length
* 215921 223332: contig of 13412 bp in length
* 223333 229433: gap of unknown length
* 229433 230858: contig of 1426 bp in length
* 230858 230959: gap of unknown length
* 230959 231959: contig of 1001 bp in length.
*
* Location/Qualifiers
*     1..231959
*         /organism="Rattus norvegicus"
*         /mol_type="genomic DNA"
*         /db_xref="taxon:10116"
*         /clones="CH230-150D22"
*     1..1341
*         /note="wgs_end_extension"
*         clone_end.T7"
*     2786..3573
*         /note="clone_boundary"
*         clone_end.T7"
*         site:
*             end sequence: RWBAX23TUB"
*         215921..216974
*         /note="wgs_contig"
*
* ORIGIN
*
* Query Match      80.7%; Score 1009.4; DB 2; Length 231959;
* Best local Similarity 91.7%; Pred. No. 4.7e-303;
* Matches 1078; Conservative
*
* QY 74 TGCATGGATGGGATCAAGAGTCAGCTCTGTACGATGGCCTACAGAGGATTT 133
* DB 145742 TGCATGGATGGGATCAAGAGAGATCTGGACGATGACGTCTACGGGGGATCC 145801
*
* QY 134 CTGGCTCTTCAACAATCCAGTATACAGCCAGGAGAACAAACGCTTCTTAAAGTTCAA 193
* DB 145802 CTGGCTCTTCAACAATCCAGTATACAGCCAGGAGAACAAACGCTTCTTAAAGTTCAA 145861
*
* QY 194 GGAGGCTTTTTCGCTGCTGTACCTGGTAGTGTGTTGCTTTGGACTGCTAGGAACTC 253
* DB 145862 GGAGGCTTTTTCGCTGCTGTACCTGGTAGTGTGTTGCTTTGGACTGCTAGGAACTC 145921
*
* QY 254 CTTGGTCTGATTATATATACATTTTCTACCAAGAGCTGAGGACTCTGACAGATGTGTTCT 313
* DB 145922 CTTGGTCTGATTATATATACATTTTCTACCAAGAGCTGAGGACTCTGACAGATGTGTTCT 145981
*
* QY 314 GCTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 373
* DB 145982 GCTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 146041
*
* QY 374 AGGCACCTATGAGTGGTCTTTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 433
* DB 146042 AGGCACCTATGAGTGGTCTTTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 146101
*
* QY 434 AATGAATCTTACGTCTCCATGCTCTCTCTCACTGATCATCAGATGATGATGATGATGATGATGAT 493
* DB 146102 AATGAATCTTACGTCTCCATGCTCTCTCTCACTGATCATCAGATGATGATGATGATGATGAT 146161
*
* QY 494 AGTGTGCCAGGTACCAAGGCTTCAACCGGAGGCTAGTGGAGATCTGGGGCCAAAGT 553

```

```

Db 146162 AGTGTGCCAGGTACCAAGGCTTCAACAGCAGGCTAAGTGGAGATCTGGGGCCAAAGT 146221
QY 554 CATTGCTTGTCTCAATTTGGTGGTCTCCCTGTTGGTCTTTTGGCCACAGATCATCTATGG 613
Db 146222 CATTGCTTGTCTCAATTTGGTGGTCTCCCTGTTGGTCTTTTGGCCACAGATCATCTACAG 146281
QY 614 CCATGCTTCAAGATATTGCAAGCTTATCTGTCAGTACCACAGTGGAGATATCCACTAT 673
Db 146282 CAGATTTGGACATATTGACGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 146341
QY 674 GGTCTTCTTATACAGTACTCTGGGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 733
Db 146342 GGTCTTCTTATACAGTACTCTGGGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 146401
QY 734 CTACTCAGGCAATATCAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
Db 146402 CTACTCAGGCAATATCAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 146461
QY 794 AAAGATCATCTTCTTGTAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853
Db 146462 AAAGATCATCTTCTTGTAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 146521
QY 854 CATGTTAATCCAAAGTACAGCTGGGATCTATACCATTAACAGCTTTAAGTATGCTAT 913
Db 146522 CATGTTAATCCAAAGTACAGCTGGGATCTATACCATTAACAGCTTTAAGTATGCTAT 146581
QY 914 CGTAGTGACAGGCTATAGCATACTTTTGGGCTTGGCTTAAACCTGCTTACTTTATGCTT 973
Db 146582 CTTAGTGACAGGCTATAGCATACTTTTGGGCTTGGCTTAAACCTGCTTACTTTATGCTT 146641
QY 974 TGTGGCTTAAAGTTCGGGAAAGCTGCGAACTTATGAAGATATCGGCTGCTCTC 1033
Db 146642 TGTGGCTTAAAGTTCGGGAAAGCTTGAAGATATCGGCTGCTCTC 146701
QY 1034 TCACCTGGGCTCTCAAGTCAATGGAAGTCTCTGAGGAGAGTTCAGAGCTTGTGCTGCT 1093
Db 146702 TCGTCTGGGCTCTCAAGTCAATGGAAGTCTCTGAGGAGAGTTCAGAGCTTGTGCTGCT 146761
QY 1094 CTCCCAATATAGACACCAAGTATGTTTCAATTTAGTAGGCTTGGCCACTTTAG 1153
Db 146762 CTCCCAATATAGACACCAAGTATGTTTCAATTTAGTAGGCTTGGCCACTTTAG 146821
QY 1154 GAAGTTAATACAAATTC-TAGGAGCATGCTGCTATCATTTGATGATGATGATGATGATGATGATGAT 1212
Db 146822 GAAGTTAATACAAATTC-TAGGAGCATGCTGCTATCATTTGATGATGATGATGATGATGATGAT 146881
QY 1213 TTTGCTTATAGCTGCTGGAGTATCATGGAAGCC 1247
Db 146882 TTTGCTTATAGCTGCTGGAGTATCATGGAAGTC 146916

RESULT 6
AC107545 254420 bp DNA linear HTG 13-MAY-2003
LOCUS Rattus norvegicus clone CH230-226016, *** SEQUENCING IN PROGRESS
DEFINITION *** 4 unordered pieces.
ACCESSION AC107545 GI:30581376
VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 254420)
AUTHORS Muzny,D.,Marle., Metzker,M.,Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Ayagi,A., Ayodeji,M., Baca,B., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Blawie,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,

```

Query Match	80.7%;	Score	1009.4;	DB 2;	Length	254420;
Best Local Similarity	91.7%;	Pred. No.	4.8e-303;			
Matches 1078;	Conservative	0;	Mismatches	96;	Indels	1; Gaps
QY	74	TGCCATGATCATGGGATCAAGAGTCAGGCTGTAGCTGTGACGTGGSCACTACGAGGAGATTT	133			
Db	199365	TGCCATGATGATGGGATCAAGAGAAAGATCTGGACGATGACGCTTACGGGGGAGATCC	199424			
QY	134	CTGGCTCTTCAACAAATTCCTAGTGATTAACGCCAGGAGAAACAAACGTTCTCTAAAGTTCAA	193			
Db	199425	CTGGCTCTTCAACAACTCCAGTGGTAAACGCCAGGAAACAAACGCTTCTCTAAAGTTCAA	199484			
QY	194	GGAGTCTTTTTCGGCTGTGTGTGTACCTGGTAGTGTGTGTCTTTGGACTGCTAGGAACTC	253			
Db	199485	GGAGGTTTTTTCGGCTCGCATATCTGTGTAGTGTGTGTCTTTGGACTGCTAGGGAACTC	199544			
QY	254	CCTGTTCTGTATTATATACATTTTCTACCAAGAGCTGAGGACCTCTCACAGATGTGTCTCT	313			
Db	199545	CCTGTTCTGTATTATATACATTTTCTACAGAGCTGAGGAGTCTGCACAGCTGTCTCT	199604			
QY	314	GCTGAACCTTGCCCTGGCTGACCTGGTGTGTGTGTGTACTGTGCCCTTTTGGCCCTATGCG	373			
Db	199605	GCTGAACCTTGCCCTGGCTGACCTGGTGTGTGTGTGTACTGTGCCCTTTTGGCCCTACGC	199664			
QY	374	AGGCACCTATGAGTGGGTCTTTTGGCACAGTCATGTGCAAACCTCTTCGAGGCATGTATAC	433			
Db	199665	AGGCACCTATGAGTGGGTCTTTTGGCACAGGCATGTGCAAACCTCTTCGAAAGATGTATGCG	199724			





```
Db 1916 AGCTTGCTCATCTGGGTGATATCCCTGCTGGTTCCTCTGCCCCCAATTAATATGGAAT 1975
QY 618 GTTCAAGATTATGCAAGCTTATCTGTGAGTACCACTGAGGAGATATCCACTATGTT 677
Db 1976 GCTTTAATCTCGAAGCTCATATGTGGTTACCATGACGAGCAATTTCCACTGTGTT 2035
QY 678 CTTGTATACAGATCACTCGGGTCTTCTCTGCAATGCTCTCACTATGATCTGTGCTAC 737
Db 2036 CTTGCCACCACTGACACTGGGGTCTTCTTGGCCACTGCTCACCATGATGTTGCTGCTAT 2095
QY 738 TCAGCAATATCAAGACCTTGCTTCTGCTGCAATCTTCCAGAGCACAATCTCTAAG 797
Db 2096 TCAGTCATAATCAAAACACTGCTTCTGCTGAGGCTTCCAGAGCAGATCTCTAAG 2155
QY 798 ATCATCTTCTTGTAGTGGCTGTGTTCTGCTGACCCAGACACCCCTTCAACCTTGCCATG 857
Db 2156 ATCATCTTCTGCTGATGCTGTGTTCTGCTGACCCAGATGCTTCAACCTCATGAAG 2215
QY 858 TTAATCCAAAGTACAAAGCTGGAGTACTATACCATCAACGAGCTTTAAGTATGCCATGTA 917
Db 2216 TTCAATCCGAGCACACTGGGAATACTATGCTCATGACGAGCTTTCACTACACCATCATG 2275
QY 918 GTGACAGAGGCTATAGCATCTTTCGGGCTTGCCTTAACTTACCTGCTTATGCTTCTTCT 977
Db 2276 GTGACAGAGGCTATGCTATCTGAGGCTGCTTAACTTACCTGCTTATGCTTCTTCTGCT 2335
QY 978 GGCCTTAAAGTCCGAAAGAACTCTGGAACCTTATGAAGGATATCGGCTGCTCTCTCAAC 1037
Db 2336 AGCCTGAAGTTTCGAAAGAACTTCTGGAACCTTGTGAAGGACATTTGCTGCTCTCTCAAC 2395
QY 1038 CTGGAGTCTCAAGTCAATGGAAGTCTCTGAGGACAGTCCAGACATTTCTGCTCTCTC 1097
Db 2396 CTTGGGCTCTCAATCAATGGAATCTTCTGAGGACATTTCAAGACATTTTCTGCTCTCTC 2455
QY 1098 CACAATGTAGACACACCACTATGTTTCAATTTAGTAGGCTTCCACACCTTACGAGAG 1157
Db 2456 CACAATGTGGAGCCACCACTATGTTTCAATTTAGTAGGCTTCCACACCTTACGAGAG 2512
QY 1158 TTAATAACAGATTTCTAGAGCATGCTGTATCATTTGGATGACACAGAAAGCTTTGCT 1217
Db 2513 ACTGCTCTGGAATTTGCAAGGATGCTGTGCTCTTCTGATGTGTGTGAGGAGGCTTTGT 2572
QY 1218 TTATAGCATGTGGAGTATCATGAGAG 1245
Db 2573 TTATAGCTTGCATCTCATGAGAG 2600

RESULT 8
BC033584
LOCUS
DEFINITION Homo sapiens chemokine (C-X-C motif) receptor 5, mRNA (cdna clone
MGC:34420 IMAGE:5225727), complete cds.
ACCESSION BC033584
VERSION BC033584.1 GI:21707372
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1923)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,I., Wang,J., Hsieh,F.,
Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Donald,M.F., Casavant,T.L.,
Schaefer,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.W.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
```

```
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Buffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Duckworth,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzyzinski,M.I., Skalska,U., Smalilus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)
22389257
PUBMED 12477932
2 (bases 1 to 1923)
Strausberg,R.
Direct Submission
Submitted (02-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcapps-email.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@hri.nih.gov
Akhter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maekeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 49 Row: n column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5730105.
FEATURES
Location/Qualifiers
source
1..1923
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:34420 IMAGE:5225727"
/tissue_type="Pancreas, Spleen, adult pooled"
/clone_lib="NIH MGC 120"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
gene
1..1923
/genes="CXCR6"
/notes="synonyms: STRL33, TYMSTR, BONZO"
/db_xref="LocusID:10663"
/db_xref="MIM:605163"
33..1061
/codon_start=1
/product="G protein-coupled receptor TYMSTR"
/protein_id="AAH33584.1"
/db_xref="GI:21707373"
/db_xref="LocusID:10663"
/translators="MAEDHYDHYGFSSFNDSQEEHQDLPQSFVFLPCMLVWFVC
GLVNSILVLSIFVHKLOSLTDVFLVNLPLADLVVCTLPFWAYAGIHEWFGVMC
KSLIGYITNFYSMLILCITVDRTIVVVKATKAYNQAKENTGKVTSLLLVLSL
LVSLPQIYGNVFNLDKLCIGYDEALSTVLAQTQTLGFLFLINIVCYSVIKTL
ILHAGQKRSKLIIFLVNAVFLTLQPNLMKFIKSTHWYIYATMSFHTINVTAI
AYIRACINPVLYAFVSLKFRKFWKLVDGICLLPYLGVSQWKSSEDNSKTFASHNV
EATSMFQL"
174..896
/misc_feature
/notes="7tm_1; Region: 7 transmembrane receptor (rhodopsin
family)"
```

ORIGIN		/db_xref="CDD:pfam00001"		Db		1026 CACAAATGCGAGCCACCAGCATGTTCCA---GTTATAGGCTTGGCAGGTTTCGAGAA		1082	
Query Match		56.5%; Score 706.4; DB 9; Length 1923;		QY		1158 TTAATAACAGAAATTCAGGAGCATGGCTGTATCATTTGGATGCAACAGAAAAGCTTTGC		1217	
Best Local Similarity		78.6%; Pred. No. 1.1e-208;		Db		1083 GCTGCTCGGAATTTGCAAGGCATGGCTGTGCCCTCTTGATGTGGTGGAGGCTTTGT		1142	
Matches 871; Conservative 0; Mismatches 231; Indels 6; Gaps 2;				QY		1218 TTATAGCATGTGAGTATCATGGAGAAG		1245	
				Db		1143 TTATAGCTTGGCATCTCTATGGAGAAG		1170	
RESULT 9				AR049107		2238 bp		DNA	
LOCUS		Sequence 1 from patent US 5824504.		AR049107		linear		PAT 29-SEP-1999	
DEFINITION				AR049107					
ACCESSION		AR049107.1		GI:6005146					
VERSION									
KEYWORDS		Unknown.							
SOURCE		Unknown.							
ORGANISM		Unclassified.							
REFERENCE		1 (bases 1 to 2238)							
AUTHORS		Elshourbady, N.A., Bergsma, D.J.P.O.Box.1.5.3.9. and Ellis, C.E.							
TITLE		Human 7-transmembrane receptor and DNA							
JOURNAL		Patent: US 5824504-A 1 20-OCT-1998;							
FEATURES		Location/Qualifiers							
source		1..2238							
		/organism="unknown"							
		/mol_type="unassigned DNA"							
ORIGIN									
Query Match		56.5%; Score 706.4; DB 6; Length 2238;		QY		141 TTCACAAATTCAGTGATAACAGCC--AGGAGAACAAACGCTTCCTAAAGCTTCAAGGAG		197	
Best Local Similarity		78.6%; Pred. No. 1.2e-208;		Db		365 TTCAGAGTTTCAATGACAGCAGCAGGAGGAGCATCAAGACTTCCTGCAGTTCAGCAAG		424	
Matches 871; Conservative 0; Mismatches 231; Indels 6; Gaps 2;				QY		198 GTCTTTTGGCCCTGTGTACTGCTGGTAGTGTGTTGTCTTTGGACTGCTAGGAAACTCCCTG		257	
				Db		425 GTCTTTTGGCCCTGTGTACTGCTGGTAGTGTGTTGTCTTTGGACTGCTAGGAAACTCCCTG		484	
				QY		258 GTCTGATATATACATTTCTACAGAGCTGAGACTCTGACAGATGCTGTTCTGCTG		317	
				Db		485 GTGCTGGTCAATCCATCTTACCATAAAGTTGAGAGCCTGACGGATGTTCTCTGGTG		544	
				QY		318 AACTTGGCCCTGGCTGACCTGGTGTGTTGTTCTGTACTCTGCCCTTTTGGCCCTATGCAGGC		377	
				Db		545 AACCTACCCCTGGCTGACCTGGTGTGTTGTTGCTGCACTCTGCCCTTCTGGCCCTATGCAGGC		604	
				QY		378 ACCTATGAGTGGTCTTTGGCAGAGTCAATGTGAAAACCTTTCGAGGCAATGATACAAATG		437	
				Db		605 ATCCATGAATGGTGTGTTGGCAGGTCAATGTGCAAGAGCCTACTGGGCATCTACATATT		664	
				QY		438 AACTTCTACGTGCTCATCTCACTCCCTGCACTACAGTGGATCGTTTCATTGTAGTG		497	
				Db		665 AACTTCTACGTGCTCATCTCACTCCCTGCACTACAGTGGATCGTTTCATTGTAGTG		724	
				QY		498 GTCCAGGCTACCAAGCCCTTCAACCGGAGGCTCAAGTGGAAAGATCTGGGGCAAGTCAATT		557	
				Db		725 GTTAAGGCCCAAGGCCCTACCAACAGCAAGAGGATGACCTGGGGCAAGGTCAQC		784	
				QY		558 TGCTTGTCTCAATTCGGTGGTCTCCCTGTGGTGTCTTGGCAGAGATCATCTATGGCCAT		617	
				Db		785 AGCTTGTCTCAATTCGGTGGTGTCTCCCTGTGGTGTCTTGGCAGAGATCATCTATGGCCAT		844	
				QY		618 GTTCAAGATATGCAAGCTTATCTGTGACGTACCAAGTGGAGAGATATCCACTATGTT		677	
				Db		845 GTCTTTAATCTCGAACAAGCTCATATGTGGTTTACCATGACGAGCAATTCACCTGTGGTT		904	
				QY		678 CTTGTTATACAGATGACTCTGGGTTCTTCTGCGCATGCTCCTACTATGATCTGTGCTAC		737	

905	DB	CTTGCCACCCAGATGACACTGGGGTCTTCTTGGCACTGCTCACCATGATTGTCTGCTAT	964
738	QY	TCAGGCATTATCAAGACCTTGTCTTTCATGCTCGAAACTTCCAGAAGCACAATCTCTAAAG	797
965	DB	TCAGTCATAATCAAAACACTGCTTTCATGCTGGAGCCTCCAGAAGCACAAGATCTCTAAAG	1024
798	QY	ATCATCTTCTTGTTAGTGGCTGTGTTCTGTGTGACCCAGACACCGCTTCAACCTTGCCTATG	857
1025	DB	ATCATCTTCTGGTGTATGGCTGTGTTCTGTGTGACCCAGATGCCCTTCAACCTCATGAAG	1084
858	QY	TTAATCCAAAGTACAAGCTGGGAGTACTATACCATAAACCGACTTTAAGTATGCCATCGTA	917
1085	DB	TTCAATCCGACACACACTGGGAATACTATGCCATGACCCAGCTTTCACACACCATCATG	1144
918	QY	GTGACAGAGGCTATAGCATACTTTCCGGGCTTGCCTTAACCGTGTACTTTATGCCCTTTGTT	977
1145	DB	GTGACAGAGGCCATCGCATACCTGAGGCGCTGCCTTAACCGTGTCTATGCCCTTTGTC	1204
978	QY	GGCTTAAAGTTCGGGAAGAACGTCGTGAAACTTATGAAGGATATCGCGTGCCTCTCTCAC	1037
1205	DB	AGCCTGAAGTTTCGAAAGAACTTCTGGAACCTTGTGAAGGACATGGTTCCTTAC	1264
1038	QY	CTGGGAGTCTCAAGTCAATGGAACTCTTCTGAGACAGTTCCAAGACTTGTTCGCGCTCC	1097
1265	DB	CTTGGGCTCTCAATCAATGAAATCTTCTGAGACAAATCCAAGACTTTTTCGCGCTCC	1324
1098	QY	CACAATGTAGACACACCAGTATGTTCCAAATGTGTAGTAGGCGCTTGCACACTTAGAGAAG	1157
1325	DB	CACAATGTGAGGCCACACGATGTTCCA--GTTATAGGCGTTGCCAGGGTTTCGAGAA	1381
1158	QY	TTAATAACAGATCTAGGAGCATGGGTATCATTTTGGATGCAACAAGAAAAGCTTTGC	1217
1382	DB	CTGCTCTGGAATTCGAAGCATGGGTGTCCTCTTGATGTGGTAGGACGAGCTTTGT	1441
1218	QY	TTATAGCATGTGGAGTATCATGGAGAAG	1245
1442	DB	TTATAGCTTGGCAATCTCATGGAGAAG	1469

```

RESULT 10
AX549062
LOCUS
DEFINITION
ACCESSION
SEQUENCE
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
AX549062
Sequence 347 from Patent WO02061087.
AX549062
AX549062.1 GI:25813841
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Burmer,G.C., Roush,C.L. and Brown,J.P.
Antigenic peptides, such as for G protein-coupled receptors
(GPCRs), antibodies thereto, and systems for identifying such
antigenic peptides
Patent: WO/02061087-A 347 08-AUG-2002;
Lifespan Biosciences, Inc. (US)
Location/Qualifiers
1..1953
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

```

ORIGIN		Score	DB 6;	Length
Query Match	56.4%;	705.4;	DB 6;	1953;
Best Local Similarity	78.5%;	Pred. No. 2.4e-208;		
Matches 870; Conservative	0;	Mismatches 232;	Indels 6;	Gaps 2;
QY	141	TTCAACAATTCAGTGTATTAACAGCC--AGGAGACAAACGCTTCCTAAAGTTCAGGAG	197	
Db	115	TTCAGCAGTTTCAATGACAGCAGCAGGAGGATCAAGACTTCTCGAGTTCAGCAAG	174	
QY	198	GTCTTTTTCGCCGTGTGTACTGTGGTGTGTTTGTCTTTTGGACTGTAGAAACTCCCTG	257	

RESULT 11

HSY13248      1953 bp      mRNA      linear      PRI 01-SEP-1997  
LOCUS      Homo sapiens mRNA for orphan chemokine receptor TYMSR.  
DEFINITION      Y13248  
ACCESSION      Y13248.1      GI:2370179  
VERSION      orphan chemokine receptor.  
KEYWORDS      Homo sapiens (human)  
SOURCE      Homo sapiens  
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
REFERENCE      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1      Loetscher, M., Anara, A., Oberlin, E., Brass, N., Legler, D.F.,  
AUTHORS      Loetscher, P., D'Apuzzo, M., Meese, E.U., Rousset, D., Virelizier, J.B.,  
Beggiani, M., Arenzana-Seisdedos, F. and Moser, B.  
TITLE      TYMSR, a putative chemokine receptor selectively expressed in  
activated T cells, exhibits HIV-1 coreceptor function  
JOURNAL      Unpublished  
REFERENCE      2 (bases 1 to 1953)  
AUTHORS      Loetscher, M.  
TITLE      Direct Submission  
JOURNAL      Submitted (19-MAY-1997) M. Loetscher, Theodor-Kocher Institute,  
University of Bern, Freiestrasse 1, CH-3012 Bern, SWITZERLAND  
FEATURES      Location/Qualifiers  
1..1953  
    /organism="Homo sapiens"  
    /mol\_type="mRNA"  
    /db\_xref="taxon:9606"  
    /chromosome="3"  
    /clone="MLRB2"  
    /cell\_line="KT30 (CD4+ T cell line)"  
    /cell\_type="T lymphocytes"  
    /tissue\_type="blood"  
1..1953  
    /gene="TYMSR"  
82..1110  
    /gene="TYMSR"  
    /codon\_start=1  
    /product="orphan chemokine receptor"  
    /protein\_id="CAA73698.1"  
    /db\_xref="GI:2370180"  
    /db\_xref="GOA:O00574"  
    /db\_xref="SWISS-PROT:O00574"  
    /translation="MAEHDEHYGFSSFNDSQEBHDFLOPSKVLPCMLVVFVC  
GLVNSLVIVISIFYHKLQSLDVLNPLADLVFVCTLPFWAYAGIHEWFGQVNC  
KSLGIIVINFTYMLTICITVDRIIVVKAAYNQQRMTWGWKVTSLIIVLSL  
LVSLPQIYGVNFDLICIYHDEAISTVVLATQMTLGFLLTMIVCYSLIITL  
LHAGFOKRLSLIIFLWAVFLITQMPFLMKFIRSTHWEYAMTSFHTIMVTAL  
AVLRACNPLVYAFVSLKFRKNFWKLVKDIGCLPYLGVSHQWKSSEDNSKTSASHV  
EATSMFQL"  
1953  
    /polyA\_site  
    /gene="TYMSR"

378      ACCTATGAGTGGGCTTTTGGCAGTCATGTGCAAAACTCTTCGAGGCATGTATACATG 437  
Db      |||  
355      ATCCATGAATGGTGTGGCCAGTCTATGTGCAAGAGCCTACTTGGGCATCTACATTT 414  
Qy      |||  
438      AACTTCTAGTGTCTCAGTCTCACTCTCACTGATCAGTGGATCGTTTCATTTAGTGTG 497  
Db      |||  
415      AACTTCTACAGTCCATGCTCATCTCACTGATCAGTGGATCGTTTCATTTAGTGTG 474  
Qy      |||  
498      GTCCAGGCTACCAAGGCTTCAACCGGCGAGCTAAGTGGAGATCTCGGCGCAAGTCATT 557  
Db      |||  
475      GTTAAGGCCCAAGGCTTACCAACGACCAAGAGGATGACCTGGGCGAAGGTCAACC 534  
Qy      |||  
558      TGTCTGCTATTTGGGTGGTCTCCCTGTGTGTCTTTCGACAGATCATCTATGGCCAT 617  
Db      |||  
535      AGCTTGTCTCATCTGGGTGATATCCCTGTGTGTCTTTCGCCCCAAATATCTATGGCAAT 594  
Qy      |||  
618      GTTCAAGATATTCAAGCTTATCTGTCAGTACCAAGTGGAGATATCCATCATGTT 677  
Db      |||  
595      GTCTTAAATCTCGACAGCTCATATGTGTACCATGACGAGCAATTTCCACATGTGTT 654  
Qy      |||  
678      CTGTGTTATACAGATGACTCTGGGTCTTCTCTGCGATGCTCATCTATGATCTGTGTAC 737  
Db      |||  
655      CTTCGCCACCATGACACTGGGTCTTCTTGTGCACTGCTCACCATGATTTGTCTGTAT 714  
Qy      |||  
738      TCAGGCATTATCAAGACCTTCTTCATGCTCGAACTTCAGAAAGCAAAATCTCTAAAG 797  
Db      |||  
715      TCAGTCATATCAAAACACTGCTTCATGCTGAGGCTTCAGAGACACAGATCTCTAAAG 774  
Qy      |||  
798      ATCATCTTCTCTTGTAGTGGTGTCTGCTGACCCAGACACCCCTTCAACCTTCCCATG 857  
Db      |||  
775      ATCATCTTCTCTGTTGATGGTGTCTGCTGCTGACCCAGATGCGCTTCAACCTCATGAAG 834  
Qy      |||  
858      TTAATCCAAAGTACAGCTGGAGTACTATACCATACCAAGCTTTAAGTATGCCATCGTA 917  
Db      |||  
835      TTCAATCCGACACACACTGGGAATATATGCCATGACCACTTTTACTACCATCATG 894  
Qy      |||  
918      GTGACAGAGGCTATAGCATACTTTTGGGCTTGCCTTAAACCTTGACTTTATGCTTTGTT 977  
Db      |||  
895      GTGACAGAGGCTATGCTATCTGAGGCGCTTAAACCTTGCTCTATGCTCTTGTGTC 954  
Qy      |||  
978      GCGCTAAAGTTCCGGAAGAAAGCTCTGGAACCTTATGAGGATATCGCTGCTCTCTAC 1037  
Db      |||  
955      AGCCTGAAGTTTCGAAAGAACTTCTGGAACCTTGTGAAGGACATTTGCTCTCTTAC 1014  
Qy      |||  
1038      CTGGGAGTCTCAAGTCAATGAAGTCTTCTGAGACAGTTCAGAGACTTGTTCGCTCC 1097  
Db      |||  
1015      CTGGGCTCTACATCAATGGAATCTTCTGAGGACATTTCAAGACTTTTCTGCTCC 1074  
Qy      |||  
1098      CACAATGTAGAGACCAAGTATGTTCCAAATGTAGTAGGCTTCCACACTTAGAGAG 1157  
Db      |||  
1075      CACAATGTGAGGCCACCAAGATGTTCCA--GTTATAGGCTTCCAGGGTTCGAGAA 1131  
Qy      |||  
1158      TTAATAACAGAAATCTAGGAGCATGCTGTATCATTTGGATGCAACAAAGAAAGCTTTC 1217  
Db      |||  
1132      GTGCTCTGGAATTTGCAAGTCAAGCTGTGCTCTTGTGATGTGGTGGAGCGCTTTGT 1191  
Qy      |||  
1218      TTATAGCATGTGGAGTATCATGGAGAAG 1245  
Db      |||  
1192      TTATAGCTTGGCAATCTCATGGAGAAG 1219

RESULT 12  
HSU73529      1918 bp      mRNA      linear      PRI 23-JUN-1997  
LOCUS      Human G protein-coupled receptor STRL33.1 (STRL33) mRNA, complete  
DEFINITION      cds.  
ACCESSION      U73529  
VERSION      U73529.1      GI:2209284  
KEYWORDS      Homo sapiens (human)  
SOURCE      Homo sapiens  
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE	1 (bases 1 to 1918)
AUTHORS	Liao,F., Alkhatib,G., Peden,K.W., Sharma,G., Berger,E.A. and Farber,J.M.
TITLE	STRL33, A novel chemokine receptor-like protein, functions as a fusion cofactor for both macrophage-tropic and T cell line-tropic HIV-1
JOURNAL	J. Exp. Med. 185 (11), 2015-2023 (1997)
MEDLINE	97311099
PubMed	9166430
REFERENCE	2 (bases 1 to 1918)
AUTHORS	Liao,F. and Farber,J.M.
TITLE	Direct Submission
JOURNAL	Submitted (07-OCT-1996) Laboratory of Clinical Investigation, National Institute of Allergy and Infectious Diseases, NIH, 9000 Rockville Pike, Bldg 10, Rm 11N-228, Bethesda, MD 20892, USA
FEATURES	Location/Qualifiers
source	1..1918
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/chromosome="3"
	/cell_type="lymphocytes"
	/note="expressed in tumor infiltrating lymphocytes"
gene	1..1918
CDS	/gene="STRL33"
	31..1059
	/note="STRL33"
	/note="contains seven transmembrane domains"
	/codon_start=1
	/product="G protein-coupled receptor STRL33.1"
	/protein_id="AA61456.1"
	/db_xref="GI:2209285"
	/translation="MAEDYHEDYGFSDNSOEQDFLQSKVFLPCWLVVWVFC KLSGLIVLVISIFPVKVLQSLTVDFLVNLPLADLVFVCTLPFWAYAGIHSMVGYQVLC GSLGIVTINPTYSMLILTCITVDVDFVWVKAYKANOQAKRMTWGWTSLLIIVLIL LVSLPQIIYGNVFNLDKILGGRDEALISTVLATQMTLGFPELLIMIVCTIVLITKAI LHAGGFQKRSILKZILFVMAVFLTQPFNLMPFIRSTHWEIVYAMTSFHTIMVTBAI AYLRACLNPVLVYAFVSLKFRKNFKVLKIDICGULPYLGVSQHWKSSDENSNTFSASHNV EATSMFQL"
ORIGIN	
Query Match	56.3%; Score 704.8; DB 9; Length 1918;
Best Local Similarity	78.5%; Pred. No. 3.6e-208;
Matches	870; Conservative 0; Mismatches 232; Indels 6; Gaps 2;
Qy	141 TTCACAAATTCACGTGATAACAGCC---AGGAGAACAAACGCTCTCTTAAAGTTCAAGGAG 197
Db	64 TTCAGCAGTTTCATGACAGCAGCAGCAGGAGCAGCATCAAGACTTCTCGCAGTTCAGCAG 123
Qy	198 GTCTTTTTCGCCCTGTGTGACCTGGTAGTGTGTTGTCTTTTGACCTGTAGAAACTCCCTG 257
Db	124 GTCTTTCTGCCCTGCATGACTTACCTGGTGGTGTGTCTGTGGTCTGGTGGGGAACCTCTCTG 183
Qy	258 GTTCTGATTATATACATTTTCTACCAAGCTGAGGACTCTGACAGATGTGTTCTTCTGCTG 317
Db	184 GTGCTGTGTCATATCCATCTTCTACCAATAGTTGACAGCCTGACGATGTGTTCTTGGTG 243
Qy	318 AACTTGCCCTGGCTGACCTGGTGTGTTGTCGTACTCTGCCCCCTTTGGGCGCTATGACGGC 377
Db	244 AACCTACCCCTGGCTGACCTGGTGTGTTGTGTGCACCTCTGCCCTTCTGGGCGCTATGACGGC 303
Qy	378 ACCTATCAGTGGGTCTTTGGCAGACGATCATGTGCAAAACCTCTTCGAGCGCATGTATACAATG 437
Db	304 ATCCATGAATGGGTGTTTGCCAGGTCATGTGCAAGAGCCTACTGGGCATCTACACTATT 363
Qy	438 AACTTCTACGTGTCCATGCTCATCTTCACTGATCAGCTGATCGTTTCATTGTAGTG 497
Db	364 AACTTCTACAGTGCATGCTCATCTTCACTGATGATGATGATGATGATGATGATGATGATGATG 423
Qy	498 GTCAGGCTACCAAGGCTTCAACCGCGCAGGCTAAGTGGAGATCTGCGGCGCAAGTCATT 557
Db	424 GTTAAAGCCACCAAGGCTTCAACCGCGCAGGATGATGATGATGATGATGATGATGATGATGATG 483

TITLE Direct Submission  
 JOURNAL Submitted (23-APR-2002) Genome Center, University of Washington,  
 Box 352145, Seattle, WA 98195, USA  
 COMMENT On Apr 23, 2002 this sequence version replaced gi:17027297.

----- Genome Center  
 Center: University of Washington Genome Center  
 Center Code: UMG  
 Web site: <http://www.genome.washington.edu>  
 Contact: [umgchgs@u.washington.edu](mailto:umgchgs@u.washington.edu)  
 Drafting Center: BCM  
 ----- Project Information  
 Center project name: chr-3  
 Center clone name: RP11-852E15 (bc0572)  
 ----- Summary Statistics  
 Sequencing vector: M13; L08821; 56% of reads  
 Chemistry: Dye-terminator; 44% of reads  
 Chemistry: Dye-terminator; 37% of reads  
 Chemistry: Dye-terminator; 8% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 209522 bases at least Q40  
 Consensus quality: 209746 bases at least Q30  
 Consensus quality: 209771 bases at least Q20  
 Insert size: 209772; sum-of-contigs  
 Quality coverage: 8.2x in Q20 bases; sum-of-contigs

-----  
 Overlapping Sequences:  
 5': RP11-165I16 (UMGC:bc0278) AC005669  
 3': RP11-91B8 (UMGC:bc0216) AC026349

#### Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

#### Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

ECORI				BglII				HindIII			
SeqDerMap	EngrPnt	SeqDerMap	EngrPnt	SeqDerMap	EngrPnt	SeqDerMap	EngrPnt	SeqDerMap	EngrPnt	SeqDerMap	EngrPnt
8696	8835	15519	16344	2966	2966						
6	<800	2067	2076	6382	6486						
14153	14070	9563	9576	512	<800						
3038	3060	1858	1823	449	<800						

6848	6876	1282	1264	1805	1734
440	<800	462	<800	3678	3815
7429	7388	4203	4247	437	<800
5315	5487	199	<800	1567	1549
3186	3234	4445	4430	346	<800
6959	6876	16180	16344	1438	1415
3159	3149	8783	8725	3415	3413
2762	2727	3192	3400	9698	9638
5886	5841	6244	6165	1543	1549
1745	1725	604	<800	3378	3413
1551	1615	303	<800	2332	2349
1871	1901	7626	7540	520	<800
1186	1181	2931	2969	1753	1734
5076	5089	7121	7139	4580	4573
10486	10324	919	827	3707	3815
646	<800	1495	1472	2852	2843
8906	8835	3686	3694	5743	5750
199	<800	2174	2250	2874	2862
237	<800	2842	2969	2360	2196
7120	7162	1800	1823	2376	2543
7093	7162	791	<800	651	<800
11716	11593	2269	2250	2826	2862
1643	1615	2567	2620	1219	1209
2201	2206	560	<800	2061	2043
7707	7749	1048	1051	1601	1549
3648	3594	5744	5750	6291	6322
7397	7388	9748	9576	12244	12240
16603	16829	4114	4052	6104	6091
1959	1901	3370	3400	3840	3815
414	<800	2921	2969	31	<800
1680	1615	196	<800	4833	4759
754	766	2080	2076	2973	2960
1083	1084	1600	1596	9069	9129
5961	5841	3638	3694	2030	2043
10484	10324	374	<800	2041	2043
5470	5487	4256	4247	13	<800
5666	5841	11051	10673	2010	2043

-----	5185	5315	3021	3222	5978	6091	-----
-----	130	<800	1804	1823	1108	1110	-----
-----	2174	2206	16809	16344	13385	13422	-----
-----	613	<800	4011	4052	1572	1549	-----
-----	279	<800	658	<800	2128	2196	-----
-----	4723	4676	1505	1472	265	<800	-----
-----	45	<800	10543	10224	1011	1014	-----
-----	6943	6876	7521	7540	415	<800	-----
-----			59	<800	4801	4759	-----
-----			4550	4430	5104	4953	-----
-----			110	<800	829	862	-----
-----			6055	5997	3842	3815	-----
-----					3117	3089	-----
-----					2125	2196	-----
-----					8046	7986	-----
-----					4498	4461	-----

Query Match 56.3%; Score 704.8; DB 9; Length 209772;  
 Best Local Similarity 78.5%; Pred No. 6.6e-208;  
 Matches 870; Conservative 0; Mismatches 233; Indels 6; Gaps 2;

QY	141	TTCAACAATTCAGTGATAACAGCC---AGGAGAACAAACGCTTCTCTAAAGTTCAAGAG	197
Db	132440	TTACAGAGTTTCAATGACAGCAGCCAGGAGGAGCATCAAGACTTCTCGACTTCAGCAAG	132499
QY	198	GTCCTTTTGGCCCTGTGCTACTGCTAGTGTGCTTCTTGGACTGCTAGGAATCCCTG	257
Db	132500	GTCCTTTTGGCCCTGTGCTAGTGTGCTTCTTGGACTGCTAGGAATCCCTG	132559
QY	258	GTTCTGATTATATACATTTTCTACAGAAAGTGGAGCTCTGACAGATGCTTCTCTGCTG	317
Db	132560	GTGCTGGTCATATCATCTTCTACCATAGTTGCAGAGCCTGACGGATGCTTCTCTGCTG	132619
QY	318	AACCTGCCCCTGGCTGACCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	377
Db	132620	AACCTGCCCCTGGCTGACCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	132679
QY	378	ACCTATGATGGTGTCTTTGGCAGCTGATGTGCAAACTCTTCGAGGCACTATACAATG	437
Db	132680	ATCCATGATGGTGTCTTTGGCAGCTGATGTGCAAACTCTTCGAGGCACTATACAATG	132739
QY	438	AACCTTCTACGTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	497
Db	132740	AACCTTCTACGTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	132799
QY	498	GTCAGGCTACCAAGCCCTTCAACCGCAGCTAAGTGAAGATCTGGGGCCAGTCAAT	557
Db	132800	GTTAAGGCCACCAAGCCCTTCAACCGCAGCTAAGTGAAGATCTGGGGCCAGTCAAT	132859
QY	558	TGCTTGTCTCATTTGGTGGTGTCTCCCTGTTGTTTCTTTGCCACAGATCATCTATGGCCAT	617
Db	132860	AGCTTGTCTCATTTGGTGGTGTCTCCCTGTTGTTTCTTTGCCACAGATCATCTATGGCCAT	132919
QY	618	GTTCAAGATATGACAAGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	677
Db	132920	GTTCTTAAATCTCGACAAGCTCATATGTTGTTTACCATGACGAGCAATTTCCACTGTGTT	132979

QY	678	CTTGTATACAGATGACTCTGGGGTTCTTCTGCCATTCCTCACTATGATCTGTGCTAC	737
Db	132980	CTTGTATACAGATGACTCTGGGGTTCTTCTGCCATTCCTCACTATGATCTGTGCTAT	133039
QY	738	TCAGGCATATCAAGACCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	797
Db	133040	TCAGGCATATCAAGACCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	133099
QY	798	ATCATCTTCTTGTAGTGGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	857
Db	133100	ATCATCTTCTTGTAGTGGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	133159
QY	858	TTAATCCAAAGTCAAGCTGGGAGTACTATACCAACACAGCTTTAAGTATGCCATCGTA	917
Db	133160	TTAATCCAAAGTCAAGCTGGGAGTACTATACCAACACAGCTTTAAGTATGCCATCGTA	133219
QY	918	GTACAGAGCTATAGCATACTTTTCGGGCTTGCCTTAAACCTGCTACTTTATGCTTTGTT	977
Db	133220	GTACAGAGCTATAGCATACTTTTCGGGCTTGCCTTAAACCTGCTACTTTATGCTTTGTT	133279
QY	978	GGCTTAAAGTTCGGAGAGCTCTGAACTTATGAAAGTATCGGCTGCTCTCAC	1037
Db	133280	GGCTTAAAGTTCGGAGAGCTCTGAACTTATGAAAGTATCGGCTGCTCTCAC	133339
QY	1038	CTGGGACTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTTCAAGACTTGTTCGCTCC	1097
Db	133340	CTGGGACTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTTCAAGACTTGTTCGCTCC	133399
QY	1098	CACAATGTAGACACACAGTATGTTCCAAATTTAGTATGCTTGGCAGCTTTAGAGAAG	1157
Db	133400	CACAATGTAGACACACAGTATGTTCCAAATTTAGTATGCTTGGCAGCTTTAGAGAAG	133456
QY	1158	TTAATAACAGAACTTAGGAGCATGCTGTATCATTTGGATGCAACAAGAAAGCTTTGC	1217
Db	133457	CTGCTCTGGAATTTGCAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	133516
QY	1218	TTATAGCATGTGAGTATCATGGAAG	1245
Db	133517	TTATAGCATGTGAGTATCATGGAAG	133544

RESULT 14  
 HSA312687 219553 bp DNA linear HTG 15-MAY-2002  
 Homo sapiens chromosome 3 clone RP6-153f14 map 3p21.3, \*\*\*  
 SEQUENCING IN PROGRESS \*\*\* 23 ordered pieces.

ACCESSION AJ312687  
 VERSION AJ312687.1 GI:13559234  
 KEYWORDS HTG; HTGS\_PHASE2.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
 AUTHORS Kiss, H., Yang, Y., Kiss, C., Andersson, K., Klein, G., Imreh, S. and  
 Dumanski, J.P.  
 TITLE The transcriptional map of the common eliminated region 1 (C3CER1)  
 in 3p21.3  
 JOURNAL Eur. J. Hum. Genet. 10 (1), 52-61 (2002)  
 MEDLINE 21906202  
 PUBMED 11896456  
 REFERENCE 2 (bases 1 to 219553)  
 AUTHORS Kiss, H.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-APR-2001) Kiss H., Microbiology and Tumorbiology  
 Center (MTC), Karolinska Institute, Box 280, Stockholm, S-17177,  
 SWEDEN  
 COMMENT The sequence is a consensus sequence of clone RP6-153f14 (1-148064  
 bp),  
 clone RP6-146el (47262-219553 bp), clone RP6-189g11 (partially,  
 150311-219553 bp) and clone RP4-787c23 (partially, 212754-219553  
 bp). The  
 sequencing contigs are in order and the gaps between them are

represented  
 by 100 Ns. Contig 1: 1-4409 bp Contig 2: 4510-5012 bp Contig 3:  
 5113-9787 bp Contig 4: 9888-11842 bp Contig 5: 11943-14635 bp  
 Contig 6:  
 14736-25152 bp Contig 7: 25253-34186 bp Contig 8: 34287-37648 bp  
 Contig  
 9: 37749-39128 bp Contig 10: 39229-56598 bp Contig 11:  
 56699-62191 bp  
 Contig 12: 62292-71083 bp Contig 13: 71184-81023 bp Contig 14:  
 81124-110809 bp Contig 15: 110910-113803 bp Contig 16:  
 113904-131104 bp  
 Contig 17: 131205-134372 bp Contig 18: 134473-136524 bp Contig  
 19:  
 136625-140189 bp Contig 20: 140290-148706 bp Contig 21:  
 148807-149537 bp  
 Contig 22: 149638-169311 bp Contig 23: 169412-219553 bp.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 23 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* been provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 4409: contig of 4409 bp in length  
 \* 4410 4509: gap of 100 bp  
 \* 4510 5012: contig of 503 bp in length  
 \* 5013 5112: gap of 100 bp  
 \* 5113 9787: contig of 4675 bp in length  
 \* 9788 9887: gap of 100 bp  
 \* 9888 11842: contig of 1955 bp in length  
 \* 11843 11943: gap of 100 bp  
 \* 11944 14635: contig of 2693 bp in length  
 \* 14636 14736: gap of 100 bp  
 \* 14737 25152: contig of 10417 bp in length  
 \* 25153 25252: gap of 100 bp  
 \* 25253 34186: contig of 8934 bp in length  
 \* 34187 34287: gap of 100 bp  
 \* 34288 37648: contig of 3362 bp in length  
 \* 37649 37749: gap of 100 bp  
 \* 37750 39128: contig of 1380 bp in length  
 \* 39129 39228: gap of 100 bp  
 \* 39229 56598: contig of 17370 bp in length  
 \* 56599 56699: gap of 100 bp  
 \* 56700 62191: contig of 5493 bp in length  
 \* 62192 62292: gap of 100 bp  
 \* 62293 71083: contig of 8792 bp in length  
 \* 71084 81023: contig of 9840 bp in length  
 \* 81024 81123: gap of 100 bp  
 \* 81124 110809: contig of 29686 bp in length  
 \* 110810 110909: gap of 100 bp  
 \* 110910 113803: contig of 2894 bp in length  
 \* 113804 113903: gap of 100 bp  
 \* 113904 131104: contig of 17201 bp in length  
 \* 131105 131204: gap of 100 bp  
 \* 131205 134372: contig of 3168 bp in length  
 \* 134373 134472: gap of 100 bp  
 \* 134473 136524: contig of 2052 bp in length  
 \* 136525 137624: gap of 1100 bp  
 \* 137625 140189: contig of 2565 bp in length  
 \* 140190 140289: gap of 100 bp  
 \* 140290 148706: contig of 8417 bp in length  
 \* 148707 148806: gap of 100 bp  
 \* 148807 149537: contig of 731 bp in length  
 \* 149538 169311: gap of 100 bp  
 \* 169312 169412: contig of 19674 bp in length  
 \* 169413 219553: gap of 100 bp  
 \* 219554 219553: contig of 50142 bp in length.  
 Location/Qualifiers  
 1..21953  
 /organism="Homo sapiens"

/mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /map="3p21.3"  
 /clone="RP6-153f14"  
 ORIGIN  
 Query Match 56.3%; Score 704.8; DB 2; Length 219553;  
 Best Local Similarity 78.5%; Pred. No. 6.6e-208;  
 Matches 870; Conservative 0; Mismatches 232; Indels 6; Gaps 2;  
 QY 141 TTCAACAATTCACGTATACAGCC---AGAGAACAAACGCTTCCTAAAGTCAAGGAG 197  
 Db TTAGCAGCTTTCAATGACAGCAGCAGAGAGATCAAGACCTTCCTGCGAGTTCAGCAAG 45907  
 QY 198 GTCTTTTGGCCCTGTGTACTCTGCTAGTGTCTTTTGGACTGCTAGGAAATCCCTG 257  
 Db GTCTTTTGGCCCTGTGTACTCTGCTAGTGTCTTTTGGACTGCTAGGAAATCCCTG 45967  
 QY 258 GTCTGATATATACATTTTACAGAGCTGAGGACTCTGACAGATGTGTTCTGCTG 317  
 Db GTGTGTGTCATATCCATCTTACCATAGTGCAGAGCCCTGACGGATGTGTTCTCTGGT 46027  
 QY 318 AACTTGGCCCTGGCTGACCTGCTGTTGTCTGTACTCTGCTGCTTTTGGGCTATGACGGC 377  
 Db AACTTGGCCCTGGCTGACCTGCTGTTGTCTGTACTCTGCTGCTTTTGGGCTATGACGGC 46087  
 QY 378 ACCTATGATGGGTCTTTGGCAGATCATGTGCAAAACTCTTTGAGGCGATGTATACATG 437  
 Db ATCCATGAATGGGTCTTTGGCAGATCATGTGCAAAACTCTTTGAGGCGATGTATACATG 46147  
 QY 438 AACTTCTACGTCTCATGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 497  
 Db AACTTCTACGTCTCATGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 46207  
 QY 498 GTCCAGGCTACCAAGGCTTCAACCGCAGGCTTAAAGTGAAGATCTGGGCGCAAGTCATT 557  
 Db GTTAAGGCCCAAGGCTTCAACCGCAGGCTTAAAGTGAAGATCTGGGCGCAAGTCATT 46267  
 QY 558 TGCTTGCTCATTTGGGTGCTCTGCTGTTGTTTCTTGGCAGATCATCTATGGCCAT 617  
 Db AGCTTGCTCATTTGGGTGCTCTGCTGTTTCTTGGCAGATCATCTATGGCCAT 46327  
 QY 618 GTTCAAGATATTGACAGCTTATCTGTCAGTACACAGATGAGGAGATATCCACTATGCTT 677  
 Db GTCTTTAATCTGACAGCTTATCTGTCAGTACACAGATGAGGAGATATCCACTATGCTT 46387  
 QY 678 GTTGTATACAGATGCTCTGGGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAC 737  
 Db GTTGTATACAGATGCTCTGGGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAC 46447  
 QY 738 TCAGGCTATATCAAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 797  
 Db TCAGGCTATATCAAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 46507  
 QY 798 ATCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857  
 Db ATCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 46567  
 QY 858 TTAATCCAAAGTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 917  
 Db TTAATCCAAAGTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 46627  
 QY 918 GTGACAGAGGCTATAGCATATCTTTGGGCTTGGCTTAAACCTGTACTTATGCTTGTGCTTGT 977  
 Db GTGACAGAGGCTATAGCATATCTTTGGGCTTGGCTTAAACCTGTACTTATGCTTGTGCTTGTGCT 46687  
 QY 978 GGTCTAAAGTTCGGAAGAACTGCTGGAAGTATGAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1037  
 Db GGTCTAAAGTTCGGAAGAACTGCTGGAAGTATGAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 46747  
 QY 1038 CTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTCCTCAAGACTTGTCTTGTGCTGCT 1097



```
Db 46748 CTTGGGGTCTCACATCAATGTAATCTTCTGAGGACAATTCACAGACTTTTCTGCCTCC 46807
Qy 1098 CACRATGTAGACACACAGATATCTCCAAATGTGTAGTAGGCTTCCACACTTAGAGAAG 1157
Db 46808 CACAATGTGGAGGCCACACAGATGTTCACAA--GTTATAGGCTTCCACAGGGTTTCGAGAA 46864
Qy 1158 TTAATAACAGAAATTTAGAGCATGGCTGTATCATTTGGATGCAACAAGAAAGCTTTTGC 1217
Db 46865 GCTGCTCTGGAATTTGCAAGGATGGCTGTGCTTCTGATGTGTGAGGAGGCTTTGT 46924
Qy 1218 TTATAGCATGTGGAGTATCATGGAGAAG 1245
Db 46925 TTATAGCTTGGCATTTCTCATGGAGAAG 46952

RESULT 15
AF124380 1032 bp DNA linear PRI 02-AUG-2001
LOCUS Macaca mulatta chemokine receptor bonzo (STRL33) gene, complete
DEFINITION
ACCESSION AF124380
VERSION AF124380.1 GI:4877754
KEYWORDS
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta

REFERENCE
1 (bases 1 to 1032)
AUTHORS Margulies B.J., Hauer, D.A. and Clements, J.E.
TITLE Identification and comparison of eleven rhesus macaque chemokine
receptors
JOURNAL AIDS Res. Hum. Retroviruses 17 (10), 981-986 (2001)
MEDLINE 21354176
PUBMED 11461684
REFERENCE
2 (bases 1 to 1032)
AUTHORS Margulies, B.J., Hauer, D.A. and Clements, J.E.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1999) Comp. Med., Johns Hopkins University, 720
Rutland Ave., Baltimore, MD 21205, USA

FEATURES
source
1..1032
/organism="Macaca mulatta"
/mo_i_type="genomic DNA"
/db_xref="taxon:9544"
<1..>1032
/gene="STRL33"
<1..>1032
/gene="STRL33"
/product="chemokine receptor bonzo"
1..1032
/gene="STRL33"
/note="similar to HIV/SIV receptor"
/codon_start=1
/product="chemokine receptor bonzo"
/protein_id="AAD31419.1"
/db_xref="GI:4877755"
/translation="MAEYDHYEDDGLNSFNDSSQEBHQDPLQPRKVFPLMVLVTV
CGLVNSLVLVIFHYKLSQSLDVLNPLADLVFVCTLPFWYAGIHEWIFQVM
KTLIVGVTVNFTSMILTCITVDREIVVVKATKAYNOQAKRMWGVKVICLLIIVIS
LLVSLPQIYGNVFNLDKLIQYHDEISITVLATQMTLGFPLLAMIVCYSVIKT
LHAGGFOKHSRLKIIFLVMAVFLITQTPFNVLKIRSTHWEYVAMTSPHYTIVTEA
IAYLRACINPLVLYAFVSLKFRKFNKFLVKDIGCLPILGVSHQWKSSEDNSKTFASHN
VEATSMFQL"

Query Match 55.6%; Score 696; DB 9; Length 1032;
Best Local Similarity 81.0%; Pred. No. 1.9e-205;
Matches 823; Conservative 0; Mismatches 190; Indels 3; Gaps 1;

Qy 121 ACGAGGGAGATTCTGGCTCTCTCAACAATTCAGTGATAACAGCC---AGAGAACAAAC 177
Db 17 ACTATGAAGATGATGGGTCTCTCAACAGTTTCAATGACAGCAGCCAGGAGGAGCATCAAG 76
```

Search completed: September 29, 2004, 02:07:45  
Job time : 4916 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2004, 00:26:24 ; Search time 522 Seconds

(without alignments)  
10181.026 Million cell updates/sec

Title: US-10-603-141-1

Perfect score: 1251

Sequence: 1 gaattcgacagggcaga.....gtatcatggagaagcctgac 1251

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

1: Geneseq1980s.\*

2: Geneseq1990s.\*

3: Geneseq2000s.\*

4: Geneseq2001as.\*

5: Geneseq2001bs.\*

6: Geneseq2002s.\*

7: Geneseq2003as.\*

8: Geneseq2003bs.\*

9: Geneseq2003cs.\*

10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1210	96.7	1475	2 AAV43795	Aav43795 Rodent ch
2	1053	84.2	1053	2 AAX86033	Aax86033 Membrane
3	1053	84.2	1053	2 AAX86032	Aax86032 Nucleic a
4	706.4	56.5	2238	2 AAV24017	Av24017 Human HBM
5	705.4	56.4	1953	7 ABZ42779	Abz42779 Human SIV
6	704.8	56.3	1918	2 AAV68515	Av68515 STRL33 co
7	689.4	55.1	1029	5 AAX07291	Aax07291 Pig-taile
8	689.4	55.1	1029	5 AAS07616	Aas07616 Pig-taile
9	689.4	55.1	1029	6 AEN84641	Abn84641 Pig-taile
10	685	54.8	1037	2 AAX07290	Aax07290 African g
11	685	54.8	1037	5 AAS07615	Aas07615 Green mon
12	685	54.8	1037	6 AEN84640	Abn84640 African g
13	680.8	54.4	1029	6 ABS73339	Ab73339 CDNA enco
14	679.2	54.3	1029	4 AAF85364	Aaf85364 Nucleotid
15	679.2	54.3	1029	9 ADE331648	Ad31648 Human 619
16	679.2	54.3	1071	5 AAH77725	Aah77725 Nucleotid
17	676.8	54.1	1026	2 AAZ28362	Aaz28362 Placetelet
18	676	54.0	1029	6 AAS73391	Ab73391 DNA enco
19	648	51.8	1026	2 AAX07289	Aax07289 Human HIV
20	648	51.8	1026	5 AAS07614	Aas07614 Human CDN
21	648	51.8	1026	6 AEN84639	Abn84639 Human G-p
22	471.8	37.7	581	2 AAV43794	Aav43794 Rodent ch
23	183.4	14.7	1074	3 AAA30596	Aaa30596 Human G p

24	183.4	14.7	1074	9 ADC22542	Adc22542 Human G p
25	183.4	14.7	2462	7 ABZ42725	Abz42725 Human C-C
26	183.4	14.7	2577	3 AAA88560	Aaa88560 Human CC
27	183.4	14.7	2577	6 AAD32916	Aad32916 Homo sapi
28	183.4	14.7	2577	7 ACA56497	Acas6497 Human sig
29	183.4	14.7	2577	7 ABX15306	Abx15306 Human CC
30	181.2	14.5	2751	2 AAQ66164	Aaq66164 Seven tra
31	181.2	14.5	2751	3 AAU18351	Aau18351 Murine V3
32	181.2	14.5	2751	3 AAA91711	Aaa91711 Murine 7T
33	181.2	14.5	2751	6 ABX54240	Abx54240 Mouse 7 t
34	180.2	14.4	1110	5 ABI97980	Abi97980 Non-endog
35	178.6	14.3	1074	3 AAA30715	Aaa30715 DNA enco
36	178.6	14.3	1074	9 ADC22694	Adc22694 Human G p
37	177.8	14.2	1137	7 AAD56092	Aad56092 Mouse cmk
38	177.8	14.2	1137	8 ADA02454	Ada02454 Mouse cmk
39	177.8	14.2	1137	9 ADB72193	Adb72193 Mouse cmk
40	177.8	14.2	2072	6 AAS17156	Aas17156 Mouse lym
41	177.8	14.2	2072	7 AAD56091	Aad56091 Mouse cmk
42	177.8	14.2	2072	8 ADA02453	Ada02453 Mouse cmk
43	177.8	14.2	2072	9 ADB72192	Adb72192 Mouse cmk
44	177.8	14.2	2084	6 ABK54274	Abk54274 Mouse 7 t
45	177.8	14.2	2085	3 AAA91745	Aaa91745 TIM recep

ALIGNMENTS

RESULT 1

AAV43795

ID AAV43795 standard; cDNA; 1475 BP.

AC AAV43795;

XX

DT 20-OCT-1998 (first entry)

XX

DE Rodent chemokine receptor 941D12 nucleotide sequence.

XX Chemokine; primate; human; rodent; chemokine receptor; asthma;

KW inflammatory response; immune response; leukocyte migration; GPCR;

KW leukocyte adhesion; chemottractant; modulation; antiviral response;

KW cellular morphology modification response; G-protein coupled receptor;

KW phosphoinositide lipid turnover; abnormal proliferation; regeneration;

XX atrophy; 941D12; ss.

XX Mus sp.

OS

XX

EH Key Location/Qualifiers

FT CDS 1..975

FT /tag= a

FT /product= "chemokine receptor 941D12"

FT /tag= b

FT /note= "nucleotide designated C may be C or T"

FT /tag= c

FT /note= "nucleotides 1412 and 1422 each designated c may be A, C, G or T"

FT misc\_feature

FT unsure

XX WO9832858-A2.

XX 30-JUL-1998.

XX 22-JAN-1998; 98WO-US000902.

XX 23-JAN-1997; 97US-0036715P.

XX (SCHE ) SCHERING CORP.

XX Mattson JD, Soto-Trejo H, Hedrick JA, Gorman DM, Zlotnik A;

XX WPI; 1998-427954/36.

XX P-PSDB; AAW70001.

XX

PT Rodent and primate chemokines and chemokine receptors - useful  
 PT diagnostically and therapeutically to treat conditions associated with  
 XX abnormal physiology or development e.g. inflammatory conditions.  
 XX  
 XX Example 3; Page 97-99; 105pp; English.

CC This represents a rodent chemokine receptor 941D12 nucleotide sequence.  
 CC The invention provides novel primate and rodent chemokines and chemokine  
 CC receptors. The chemokines, receptors and binding compounds (optionally  
 CC antibodies/fragments specifically binding the chemokines) are useful  
 CC therapeutically to treat conditions associated with abnormal physiology  
 CC or development e.g. inflammatory conditions such as asthma. Chemokines  
 CC are important in immune and inflammatory responses in that they induce  
 CC leukocyte migration and adhesion. They are also chemoattractants for  
 CC several cells involved in inflammation and can induce other biological  
 CC responses e.g. modulation of second messenger levels (e.g. Ca<sup>++</sup>),  
 CC cellular morphology modification responses, phosphoinositide lipid  
 CC turnover, possible antiviral responses etc. The chemokine receptors of  
 CC the invention exhibit structural properties of G-protein coupled  
 CC receptors (GPCR), although their ligands have not yet been identified.  
 CC The chemokine and chemokine receptor polypeptides are useful to produce  
 CC ligand:receptor complexes in vivo or in assay techniques. Assays may also  
 CC involve chemical antagonists which block complex production or utilize  
 CC antagonists) can be used to modulate the physiological responses in cells  
 CC (especially neurons, macrophages or lymphocytes) to treat e.g. abnormal  
 CC proliferation, regeneration, generation and atrophy. The polypeptides are  
 CC also used to produce antibodies useful diagnostically, for drug screening  
 CC or for polypeptide purification. The polynucleotides are useful to  
 CC produce probes for detecting the polypeptides, and to isolate the  
 CC polypeptides or related sequences, especially from other species. They  
 CC also allow transformation of cells for polypeptide production

XX SQ Sequence 1475 BP; 374 A; 346 C; 312 G; 443 T; 0 U; 0 Other;

Query Match 96.7%; Score 1210; DB 2; Length 1475;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 1225; Conservative 0; Mismatches 5; Indels 2; Gaps 1;  
 QY 15 GGCAGACCTTTAGTGACACACTTCACTCTGGAAACAAAGCTACTGGGCTTCTCTCTGAT 74  
 DB 4 GGCAGACCTTTAGTGACACACTTCACTCTGGAAACAAAGCTACTGGGCTTCTCTCTGAT 63  
 QY 75 GGCATGATGATGGGCATCAAGAGTCAAGTCTCTGATGGGCACTACGAGGAGATTTC 134  
 DB 64 GCCATGATGATGGGCATCAAGAGTCAAGTCTCTGATGGGCACTACGAGGAGATTTC 123  
 QY 135 TGGCTCTTCAACAATTCAGTGATACAGCCAGGAGACAAAGCTTCCCTAAAGTTCAAG 194  
 DB 124 TGGCTCTTCAACAATTCAGTGATACAGCCAGGAGACAAAGCTTCCCTAAAGTTCAAG 183  
 QY 195 GAGGTCTTTTGGCCCTGTGTGACCTGGTAGTGTGTCTTGGACTGTAGGAACTCC 254  
 DB 184 GAGGTCTTTTGGCCCTGTGTGACCTGGTAGTGTGTCTTGGACTGTAGGAACTCC 243  
 QY 255 CTGGTCTGATATATACATTTCTACCAAGAGCTGAGGACTCTGACAGATGTTCTG 314  
 DB 244 CTGGTCTGATATATACATTTCTACCAAGAGCTGAGGACTCTGACAGATGTTCTG 303  
 QY 315 CTGAACCTTGCCCTGGCTGACCTGGTGTGTGTCTGACTCTGCCCTTTGGGCTTATGCA 374  
 DB 304 CTGAACCTTGCCCTGGCTGACCTGGTGTGTGTCTGACTCTGCCCTTTGGGCTTATGCA 363  
 QY 375 GGCACCTATGAGTGGTCTTTGGCAGATCATGTGCAAACTCTTCGAGGATGATATACA 434  
 DB 364 GGCACCTATGAGTGGTCTTTGGCAGATCATGTGCAAACTCTTCGAGGATGATATACA 423  
 QY 435 ATGAATCTTACGTGTCCATGCTCACTCTCACTGCATCAGAGTGGATGTTTCAATTGTA 494  
 DB 424 ATGAATCTTACGTGTCCATGCTCACTCTCACTGCATCAGAGTGGATGTTTCAATTGTA 483  
 QY 495 GTGGTCCAGGCTACCAAGGCTTCAACCGGAGGCTTAAGTGAAGATCTGGGCGCAAGTC 554

DB 484 GTGGTCCAGGCTACCAAGGCTTCAACCGGAGGCTAAGTGAAGATCTGGGCGCAAGTC 543  
 QY 555 ATTTGCTTGTCTCATTTGGGTGGTCTCCCTGTGTGGTTCTTTGGCCACAGATCATCTATGGC 614  
 DB 544 ATTTGCTTGTCTCATTTGGGTGGTCTCCCTGTGTGGTTCTTTGGCCACAGATCATCTATGGC 603  
 QY 615 CATGTTCAAGATATTCAGCAAGCTTATCTGTCACTACACAGTGAAGAGATATCCACTATG 674  
 DB 604 CATGTTCAAGATATTCAGCAAGCTTATCTGTCACTACACAGTGAAGAGATATCCACTATG 663  
 QY 675 GTTCTTGTATACAGATCACTCTGGGTTCTTCTGCAATGCTCACTATGATCTGTGC 734  
 DB 664 GTTCTTGTATACAGATCACTCTGGGTTCTTCTGCAATGCTCACTATGATCTGTGC 723  
 QY 735 TACTCAGCATATTCAGCAAGCTTCTCTCATCTCGAAACTTCCAGAGCAAAATCTCTA 794  
 DB 724 TACTCAGCATATTCAGCAAGCTTCTCTCATCTCGAAACTTCCAGAGCAAAATCTCTA 783  
 QY 795 AAGATCATCTTCTTGTAGTGGTGTGTCTGCTGACCCAGACACCCCTTCAACCTTGGC 854  
 DB 784 AAGATCATCTTCTTGTAGTGGTGTGTCTGCTGACCCAGACACCCCTTCAACCTTGGC 843  
 QY 855 ATGTTAATCCAAAGTACAAAGCTGGGAGTACTATACCAATACCAAGCTTTAAGTATGCCATC 914  
 DB 844 ATGTTAATCCAAAGTACAAAGCTGGGAGTACTATACCAATACCAAGCTTTAAGTATGCCATC 903  
 QY 915 GTAGTGACAGAGGCTATAGCATCTTTC - GGGCTTGCCTTAACCCGTACTTTATGCCCT 972  
 DB 904 GTAGTGACAGAGGCTATAGCATCTTTC - GGGCTTGCCTTAAACCCGTACTTTATGCCCT 963  
 QY 973 TTGTTGGCTTAAAGTTCGGAAGAACGCTCTGGAACCTTATGAAGGATATCGGCTGCCCTCT 1032  
 DB 964 TTGTTGGCTTAAAGTTCGGAAGAACGCTCTGGAACCTTATGAAGGATATCGGCTGCCCTCT 1023  
 QY 1033 CTCACCTGGGAGTCTCAAGTCAATGGAAGTCTTCTAGGACAGTTCGAGACAGTTCGAGTCTG 1092  
 DB 1024 CTCACCTGGGAGTCTCAAGTCAATGGAAGTCTTCTAGGACAGTTCGAGACAGTTCGAGTCTG 1083  
 QY 1093 CCTCCCACTATGAGACACCAAGTATGTTCCAAATGTTAGTAGGCTTCCGACACTTAG 1152  
 DB 1084 CCTCCCACTATGAGACACCAAGTATGTTCCAAATGTTAGTAGGCTTCCGACACTTAG 1143  
 QY 1153 AGAAGTTAATAACAGAAATCTAGGAGATGCTGCTGATCATTTGGATGCAACAAGAAAGC 1212  
 DB 1144 AGAAGTTAATAACAGAAATCTAGGAGATGCTGCTGATCATTTGGATGCAACAAGAAAGC 1203  
 QY 1213 TTTCTTATAGCATGTGGAGTATCATGAGAA 1244  
 DB 1204 TTTCTTATAGCATGTGGAGTATCATGAGAA 1235

RESULT 2

AAx86033/c

ID AAx86033 standard; cDNA to mRNA; 1053 BP.

XX AAx86033;

AC AAx86033;

XX 15-SEP-1999 (first entry)

XX Membrane penetrating type receptor protein Et60 nucleic acid fragment.

DE 7 times membrane-penetrating type receptor protein; ET60;

KW leukocyte function; control; drug; ss.

XX Unidentified.

OS JP1155573-A.

PN 15-JUN-1999.

PD 27-NOV-1997;

XX 97JP-00325823.

PR 27-NOV-1997;

XX 97JP-00325823.

Qy	858	TTAATCCAAAGTACAAGCTGGGAGTACTATACCAAAACAGCTTTAAAGTATGCCATCGTA	911
Db	273	TTATCCAAAGTACAAGCTGGGAGTACATACCAACAACAGCTTTAAAGTATGCCATCGTA	214
Qy	918	GTGACAGAGCGCTATAGCATACTTTTGGGCTTGCCTTAACCCCTGCTACTTTATGCCTTTGTT	977
Db	213	GTGACAGAGCGCTATAGCATACTTTTGGGCTTGCCTTAACCCCTGCTACTTTATGCCTTTGTT	154
Qy	978	GGCTTAAAGTTCGGGAAGAACGCTCTGAAAACCTTATGAAGGATATCGGCTCTCTCTCAC	1037
Db	153	GGCTTAAAGTTCGGGAAGAACGCTCTGAAAACCTTATGAAGGATATCGGCTCTCTCTCAC	94
Qy	1038	CTGGAGCTCTCAAGTCAATGGAAGTCTTTCTGAGGACAGTTCGAAGACTTGTTCGCCTCC	1097
Db	93	CTGGAGCTCTCAAGTCAATGGAAGTCTTTCTGAGGACAGTTCGAAGACTTGTTCGCCTCC	34
Qy	1098	CACAATGTAGAGACCACCACTATGTCCAAATTG	1130
Db	33	CACAATGTAGAGACCACCACTATGTCCAAATTG	1
RESULT 3			
AAAX86032			
ID	AAAX86032 standard; cDNA to mRNA; 1053 BP.		
XX	AC	AC	
XX	AC	AC	AAAX86032;
DT	15-SEP-1999 (first entry)		
XX			
XX	Nucleic acid encoding a membrane penetrating type receptor protein ET60.		
DE	7 times membrane-penetrating type receptor protein; ET60;		
KW	leukocyte function; control; drug; ss.		
OS	Unidentified.		
XX			
XX	JP11155573-A.		
PN	15-JUN-1999.		
XX			
XX	27-NOV-1997; 97JP-00325823.		
PF			
XX	27-NOV-1997; 97JP-00325823.		
PR			
XX	(ASAH ) ASahi KASEI KOGYO KK.		
XX	WPI; 1999-398070/34.		
XX	P-PSDB; AA23825.		
DR			
XX	New 7 times membrane-penetrating type receptor protein ET60 - useful for		
PT	detecting drug controlling function of leukocyte.		
PT			
XX	Claim 4; Page 18-19; 20pp; Japanese.		
PS			
XX			
XX	The present sequence encodes a 7 times membrane-penetrating type receptor		
CC	protein ET60. The 7 times membrane-penetrating type receptor protein ET60		
CC	can be used for detecting a drug controlling the function of leukocyte		
CC			
XX			
SQ	Sequence 1053 BP; 244 A; 253 C; 239 G; 317 T; 0 U; 0 Other;		
Query Match 84.2%; Score 1053; DB 2; Length 1053;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	78	ATGGATGATGGGCATCAAGACTCAGCTCTGTACGATGGGCACCTACGAGGAGATTCTCG	137
Db	1	ATGGATGATGGGCATCAAGACTCAGCTCTGTACGATGGGCACCTACGAGGAGATTCTCG	60
Qy	138	CTCTTCAACAATCCAGTGATAACAGCCAGAGAACAAACGCTTCTCTAAAGTTCAAGGAG	197
Db	61	CTCTTCAACAATCCAGTGATAACAGCCAGAGAACAAACGCTTCTCTAAAGTTCAAGGAG	120
Qy	198	GTCTTTTGGCCCTGTGTACTCGTAGTGTGTCTTTTGGACTGTGAGAAACTCCCTG	257

```
Db 121 GTCCTTTGGCCCTGGTGTACTCTGGTAGTGTGTTCTTTGGAGCTGTAGAAATCCCTG 180
Qy 258 GTTCTGATTATATACATTTTCTACAGAGCTGAGGACTCTGCAGAGTGTTGTTCTGCTG 317
Db 181 GTTCTGATTATATACATTTTCTACAGAGCTGAGGACTCTGCAGAGTGTTGTTCTGCTG 240
Qy 318 AACTTGGCCCTGGCTGAGCTGGTGTGTTGTTCTGTAAGTCTGCTGCTTTGGGCTATGAGGC 377
Db 241 AACTTGGCCCTGGCTGAGCTGGTGTGTTGTTCTGTAAGTCTGCTGCTTTGGGCTATGAGGC 300
Qy 378 ACCTATGAGTGGGCTTTGGGACAGTCAATGTGCAAACTCTGAGGACATGATACAATG 437
Db 301 ACCTATGAGTGGGCTTTGGGACAGTCAATGTGCAAACTCTGAGGACATGATACAATG 360
Qy 438 AACTTCTAGTGTCCATGCTCACTCTCACTGCAATCACTGATGATGTTTCAATGATG 497
Db 361 AACTTCTAGTGTCCATGCTCACTCTCACTGCAATCACTGATGATGTTTCAATGATG 420
Qy 498 GTCCAGGCTACCAAGGCTTCAACCGGAGGCTAAGTGGAGATCTGGGCGCAAGTCAT 557
Db 421 GTCCAGGCTACCAAGGCTTCAACCGGAGGCTAAGTGGAGATCTGGGCGCAAGTCAT 480
Qy 558 TGCTTGCTCAATTTGGGCTGCTCCCTGTTGGTTCCTTTGCGCAGATCACTATGCGCAT 617
Db 481 TGCTTGCTCAATTTGGGCTGCTCCCTGTTGGTTCCTTTGCGCAGATCACTATGCGCAT 540
Qy 618 GTTCAAGATATTGCAAGCTTATCTGTCAGTACACAGTGGAGATATCCATATGTT 677
Db 541 GTTCAAGATATTGCAAGCTTATCTGTCAGTACACAGTGGAGATATCCATATGTT 600
Qy 678 CTTGTTATACAGATGACTCTGGGCTTCTTCCGCAATGCTCACTATGATCTGCTAC 737
Db 601 CTTGTTATACAGATGACTCTGGGCTTCTTCCGCAATGCTCACTATGATCTGCTAC 660
Qy 738 TCAGGCAATTACAGACTTCTGCTTCACTGCTGCAAACTTCCAGAGCACAAATCTTAAAG 797
Db 661 TCAGGCAATTACAGACTTCTGCTTCACTGCTGCAAACTTCCAGAGCACAAATCTTAAAG 720
Qy 798 ATCATCTCTCTGCTGAGTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 857
Db 721 ATCATCTCTCTGCTGAGTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Qy 858 TTAATCCAAAGTACAAGCTGGGAGTACTATACATTAACAGCTTTAAGTATGCCATGTA 917
Db 781 TTAATCCAAAGTACAAGCTGGGAGTACTATACATTAACAGCTTTAAGTATGCCATGTA 840
Qy 918 GTCCAGAGCTATAGCATACTTTCCGGCTTCCGCTTAAACCTGCTACTTTATGCCCTTGT 977
Db 841 GTCCAGAGCTATAGCATACTTTCCGGCTTCCGCTTAAACCTGCTACTTTATGCCCTTGT 900
Qy 978 GGCTTAAAGTCCGGAAGACGCTCTGGAACCTTATGAAGGATATCGGCTGCTCTCTAC 1037
Db 901 GGCTTAAAGTCCGGAAGACGCTCTGGAACCTTATGAAGGATATCGGCTGCTCTCTAC 960
Qy 1038 CTGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTCCAGACTTGTCTGCTCTCC 1097
Db 961 CTGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTCCAGACTTGTCTGCTCTCC 1020
Qy 1098 CACAATGTAGAGACCAACCAAGTATGTTTCCAATTG 1130
Db 1021 CACAATGTAGAGACCAACCAAGTATGTTTCCAATTG 1053
```

## RESULT 4

AAV24017

ID AAV24017 standard; cDNA; 2238 BP.

XX

AC

AAV24017;

XX

DT 07-AUG-1998 (first entry)

XX

DE Human HMBU14 coding sequence.

```
XX Human; 7-transmembrane receptor; HMBU14; infection; HIV; cancer; asthma;
KW anorexia; Parkinson's disease; acute heart failure; atherosclerosis;
KW osteoporosis; angina pectoris; myocardial infarction; psychotic disorder;
KW benign prostatic hypertrophy; neurological disorder; therapy;
KW Huntington's disease; Gilles de la Tourette's syndrome; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
CDS 332..1360
/*tag= a
/product= "HMBU14"
EP834563-A2.
XX 08-APR-1998.
XX 23-SEP-1997; 97EP-00307428.
XX 26-SEP-1996; 96US-0026669P.
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX Elshourbagy NA, Bergsma DJ, Ellis CE;
XX WPI: 1998-208924/19.
XX P-PSDB; AAW54041.
XX New isolated human 7-trans-membrane receptor, HMBU14 - useful for
PT developing products for treating, e.g. infections, pain, cancers,
PT anorexia, hypotension, osteoporosis and asthma.
XX Claim 4; Fig 1; 36pp; English.
```

This sequence encodes the human HMBU14 protein of the invention. HMBU14 is a human 7-transmembrane (TM) receptor. The products can be used for treating or preventing conditions related to abnormal HMBU14 expression or activity, e.g. infections such as bacterial, fungal, protozoan and viral infections especially infections caused by HIV-1 and HIV-2, pain, cancers, anorexia, bulimia, asthma, Parkinson's disease, acute heart failure, atherosclerosis, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, myocardial infarction, ulcers, asthma, allergies, benign prostatic hypertrophy and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia or severe mental retardation, and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome. The products can also be used for diagnosing a disease or susceptibility to a disease related to the expression of HMBU14 and for identifying compounds which bind to and activate or inhibit a receptor for the polypeptide

XX Sequence 2238 BP; 593 A; 512 C; 517 G; 616 T; 0 U; 0 Other;

Query Match 56.58; Score 706.4; DB 2; Length 2238;  
Best Local Similarity 78.6%; Pred. No. 3.2e-209;  
Matches 871; Conservative 0; Mismatches 231; Indels 6; Gaps 2;

```
Qy 141 TTCACCAATTCAGTGATAACAGCC---AGGAGAACAAACGCTTCTCTAAAGTTCAAGGAG 197
Db 365 TTCAGCAGTTTCAATGACAGCAGCCAGGAGGAGCATCAAGACTTCTCGAGTTTCAGCAAG 424
Qy 198 GTCTTTTGGCCCTGTGTGTAACCTGGTAGTGTGTTGCTTTGGAGTGTAGAAATCCCTG 257
Db 425 GTCTTTTGGCCCTGTGTGTAACCTGGTAGTGTGTTGCTTTGGAGTGTAGAAATCCCTG 484
Qy 258 GTTCTGATTATATACATTTTCTACCAAGCTGAGGACTCTGACAGATGTGTTCTGCTG 317
Db 485 GTGCTGTGCATATCCATCTTCTACCATAGTTTCAGAGCTGACGATGTGTTCTGCTG 544
Qy 318 AACTTGGCCCTGGCTGACCTGGTGTGTTGTTCTGTACTCTGCTCTTTTGGGCTATGAGGC 377
Db 545 AACCTACCCCTGGCTGACCTGGTGTGTTGTTGTGCACTCTGCCCTTCTGGGCTATGAGGC 604
```

QY 378 ACCTATGAGTGGGTCTTTGGGACAGTCATGTGCAAAAACCTTTCAGGCGATGTATACAAATG 437  
Db ATCCATGAATGGGTGTTTGGCCAGGTGATGTGCAAGACCTTACTGGGCATCTACACTATT 664  
QY 438 AACTTCTACGTGTCACCTCACTCTCACCTGCATCAGTGGATCGTTTCATCTAGTAGTG 497  
Db AACTTCTACGTGTCACCTCACTCTCACCTGCATCAGTGGATCGTTTCATCTAGTAGTG 724  
QY 498 GTCCAGGCTACCAAGGCCCTTCAACCGGCGAGGCTAAGTGGAAAGATCTGGGGCCAAAGTCATT 557  
Db GTTAAGGCCACCAAGGCCCTACACCCAGCAAGCCAGAGGATGACCTGGGGCAAGGTCAAC 784  
QY 558 TCGTTGCTCATTTGGGTGGTCTCCCTGTGTCTTCTTTCGACAGATCATCTATGGCCAT 617  
Db AGCTTGTCTCATCTGGGTGATATCCCTGTCTGTCTTCTTTCGCCCCAAATATATCTATGGCAAT 844  
QY 618 GTTCAAGATATTGCAAGCTTATCTGTGATGACAGTGGAGAGATATCCACTATGGTT 677  
Db GTCTTTATCTCGACAAGCTCATATGTGTTTACCATGACGAGGCAATTTCCACTGTGTT 904  
QY 678 CTGTTTATACAGATGACTCTGGGGTCTTCTGCTGCTTCTGCTCATATGATTTCTGTGTAC 737  
Db CTGTCACCCAGATGACACTGGGGTCTTCTTGTGCTGCTTCTGCTGCTGCTGCTAT 964  
QY 738 TCAGGCATTATCAAGACCTTGTCTCATGCTCGAACTTCCAGAGGACCAAAATCTCTAAAG 797  
Db TCAGTCATAATCAAAACACTGCTTCTGCTGAGGCTTCCAGAGGACAGATCTCTAAAG 1024  
QY 798 ATCATCTTCTTGTAGTGGCTGTGTTCTGCTGTGACCCAGACACCCCTTCAACCTTGGCCATG 857  
Db ATCATCTTCTTGTGATGGCTGTGTTCTGCTGTGACCCAGATGCCCCCTTCAACCTCATGAAG 1084  
QY 858 TTAATCCAAAGTACAAAGCTGGGAGTACTATACCTAACCGCTTTAAGTATGCCATGTA 917  
Db TTCACTCCGACGACACACTGGGAATACATGCTGATGACCGAGCTTTCATCACCATCATG 1144  
QY 918 GTGACAGAGGCTATAGCATCTTTCGGGCTTTCCTTAACTTACCTCTTATGCTTTCCTTTGTT 977  
Db GTGACAGAGGCTATGCTATACCTGAGGCGCTTTCCTTAACTTACCTCTTATGCTTTCCTTTGTC 1204  
QY 978 GCTTAAAGTTCGGAAGACGCTGTGAAACCTTATGAAGATATGCGTCTGCTCTCTCAC 1037  
Db AGCCTGAAGTTCGAAAGAACTTTCGAAACCTTGTGAAGGACATGTTGCTTGCCTCTAC 1264  
QY 1038 CTGGGAGTCTCAAGTCAATGGAAGCTTCTGAGGACAGTTCGAAGACTTGTTCGCTCC 1097  
Db CTGGGAGTCTCAATGGAAGCTTCTGAGGACAGTTCGAAGACTTGTTCGCTCC 1324  
QY 1098 CACAATGTAGACACCAAGTATGTTTCCAAATGTATGATAGGCTTGCACACTTACAGAA 1157  
Db CACAATGTAGGAGCCACCAAGTATGTTTCCAAATGTATGATAGGCTTGCACACTTACAGAA 1381  
QY 1158 TTAATAACAGAACTTCTAGGACATGCTGATCATTTTGGATGCAACAAGAAAGCTTTGC 1217  
Db GTGCTCTGGAATTTGCAAGGATGCTGCTGCTTCTGATGTGTGAGGAGGCTTTGT 1441  
QY 1218 TTATAGCATGTGGAGTATCATGGAGAAG 1245  
Db TTATAGCTTGGCAATTTCTCATGGAGAAG 1469

RESULT 5  
ABZ42779  
ID ABZ42779 standard; DNA; 1953 BP.  
XX  
AC ABZ42779;  
XX  
DT 04-MAR-2003 (first entry)  
XX Human SIV/HIV receptor BONZO nucleotide SEQ ID NO:347.  
DE  
DE G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;  
KW G protein-coupled receptor modulator; antibody; immune-related disease;  
KW

growth-related disease; cell regeneration-related disease; AIDS; cancer;  
immunological-related cell proliferative disease; autoimmune disease;  
Kw Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;  
Kw osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;  
Kw graft versus host disease; Parkinson's disease; multiple sclerosis; pain;  
Kw psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;  
Kw mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;  
Kw hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;  
Kw ulcer; gene; ds.  
XX Homo sapiens.  
XX WO200261087-A2.  
XX 08-AUG-2002.  
XX 19-DEC-2001; 2001WO-US050107.  
XX 19-DEC-2000; 2000US-0257144P.  
XX (LIFE-) LIFESPAN BIOSCIENCES INC.  
XX Burmer GC, Roush CL, Brown JP;  
XX P-PSDB; ABP81931.  
XX WPI; 2003-046718/04.  
XX New isolated antigenic peptides e.g., for G protein-coupled receptors  
PT (GPCR), useful for diagnosing and designing drugs for treating conditions  
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or  
PT autoimmune diseases.  
XX Disclosure; Fig 1; 523pp; English.  
XX The present invention describes antigenic peptides (I) comprising: (a)  
CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino  
CC acids. Also described: (1) an assay for the detection of a particular G  
CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;  
CC and (2) an isolated antibody having high specificity and high affinity or  
CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in  
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an  
CC antibody against a particular GPCR, and in the production of specific  
CC antibodies. The peptides and antibodies are also useful for detecting the  
CC presence or absence of corresponding GPCR. The antigenic peptides for  
CC GPCRs and antibodies are useful for diagnosing and designing drugs for  
CC treating immune-related diseases, growth-related diseases, cell  
CC regeneration-related diseases, immunological-related cell proliferative  
CC diseases, or autoimmune diseases e.g. AIDS, Alzheimer's disease,  
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,  
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute  
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host  
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,  
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory  
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,  
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or  
CC any other disorder in which GPCRs are involved. The antibodies may be  
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode  
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the  
CC exemplification of the present invention  
XX Sequence 1953 BP; 491 A; 468 C; 455 G; 539 T; 0 U; 0 Other;  
SQ

Query Match 56.4%; Score 705.4; DB 7; Length 1953;  
Best Local Similarity 78.5%; Pred. No. 6.1e-209;  
Matches 870; Conservative 0; Mismatches 232; Indels 6; Gaps 2;  
QY 141 TTCAACAATTCAGTGATAACAGCC---AGGAGAACAAAGCTTCTCTAAAGTTCAAGGAG 197  
Db 115 TTACAGCTTTTCATGACAGCAGCCAGGAGGACATCAAGACTTCTGCGAGTTGAGCAAG 174  
QY 198 GTCTTTTTCGCTGTGTGTACCTGGTAGTGTCTTTGCTTTGAGACTGTAGAAACTCCCTG 257  
Db 175 GTCTTTTTCGCTGTGTGTACCTGGTAGTGTCTTTGCTTTGCTGTGTGGGAACTCTCTG 234

QY 258 GTTCTGATTATATATATTTCTACCAAGAGCTGAGGACTCTGACAGATGTTGTTTCTGCTG 317  
DB 235 GTGCTGGTCATATCATCTTCTACCAATAAGTTGCAGAGCCTGACGGATGTTGTTCTTGCTG 294  
QY 318 AACTTGCCCTGGCTGACCTGGTGTGTTGCTGTTACTCTGCTCTTGGGCTATGAGGC 377  
DB 295 AACCTACCCCTGGGTGACCTGGTGTGTTGCTGTTACTCTGCTCTTGGGCTATGAGGC 354  
QY 378 ACCTATGAGTGGGTCTTGGCACAGTCACTGTCGAAACTCTTCCAGGCACTATACAATG 437  
DB 355 ATCCATGAATGGTGTGTTGGCCAGTCACTGTCAGAGCCTACTGCGCATCTACACTATT 414  
QY 438 AACTTCTACGTGTCATGCTCACTCTCACTGTCGATCACTGAGTATGTTTCTATGAGTG 497  
DB 415 AACTTCTACAGTCCATGCTCATCTCACCTGCATCACTGTTGATGCTGTTTCTATTGAGTG 474  
QY 498 GTCCAGGCTACCAAGGCTTCAACCGGCGAGCTAAGTGAAGATCTGGGGCCCAAGTCATT 557  
DB 475 GTTAAGGCCACCAAGGCTCAACACAGCAAGCCAGAGATGACCTGGGCAAGGTCAACC 534  
QY 558 TGTGTTGCTCACTTGGGTGCTCCCTGTTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 617  
DB 535 AGCTTGTCTCATCTGGGTGATATCCCTGCTGTTCTTCTGCTGCTGCTGCTGCTGCTGCT 594  
QY 618 GTTCAAGATATGACAGCTTATCTGTCAGTACACAGTACAGGAGATATCCATGCTGTT 677  
DB 595 GTCTTTAATCTGCAAGCTCATATGTTGTTACCATGACGAGGCAATTTCCACTGTGGTT 654  
QY 678 CTTGTTATACAGTACTCTGGGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737  
DB 655 CTTGCCACCCAGATGACACTGGGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 714  
QY 738 TCAGGCTATACAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 797  
DB 715 TCAGTCAATAATCAAAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774  
QY 798 ATCATCTTCTGTTAGTGGTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857  
DB 775 ATCATCTTCTGTTAGTGGTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 834  
QY 858 TTAATCCAAAGTACAGCTGGGAGTACTATACCATTAACCAAGCTTTAAGTATGCTGCTGCT 917  
DB 835 TTCATCCGACGACACACTGGGAACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894  
QY 918 GTGACAGAGCTATAGCATCTTTCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 977  
DB 895 GTGACAGAGGCTATAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954  
QY 978 GGCTTAAAGTTCGGGAAGAACGCTGGAAGACTTATGAAGGATATCGGCTGCTGCTGCTGCTGCT 1037  
DB 955 AGCCTGAAGTTTCAAGAACTTCTGGAAGACTTGTGAAGGACATGTTGTTGCTGCTGCTGCTGCT 1014  
QY 1038 CTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTCAGAGACTGTTGCTGCTGCTGCTGCT 1097  
DB 1015 CTGGGCTCTCATCATCAATGGAATCTTCTGAGGACAACTTCAAGAACTTCTTCTGCTGCTGCTGCT 1074  
QY 1098 CACAATGTAGAGACCAACAGTATGTTCCAAATTTAGTAGGCTTTCGACACACTTAGAGAA 1157  
DB 1075 CACAATGTAGAGGACCAACAGTATGTTCCAAATTTAGTAGGCTTTCGACACACTTAGAGAA 1131  
QY 1158 TTAATAACAAATCTAGGAGAGTGTGCTGATCATTTGGATGCAACAAGAAAGCTTTGCT 1217  
DB 1132 GTGCTCTGGAATTTGCAAGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1191  
QY 1218 TTATAGCATGTGAGTATCATGAGAA 1245  
DB 1192 TTATAGCTTGGCATCTCATGAGAA 1219

RESULT 6  
AAV68515  
ID AAV68515 standard; DNA; 1918 BP.

XX AAV68515;  
XX AC 22-JAN-1999 (first entry)  
XX DT STRL33 coding sequence.  
XX DE STRL33; human; HIV-1; HIV related disorder; therapy; CD4; diagnosis;  
XX KW fusion accessory factor; ds.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX FT CDS 31..1059  
XX FT /\*tag= a  
XX PN WO9844098-A2.  
XX PD 08-OCT-1998.  
XX PF 31-MAR-1998; 98WO-US006517.  
XX PR 31-MAR-1997; 97US-0042880P.  
XX PA (USSH ) US SEC HEALTH & HUMAN SERVICES.  
XX PI Farber JM, Liao F, Alkhatib G, Berger EA;  
XX DR WPI; 1998-557107/47.  
XX DR P-PSDB; AAW80806.  
XX PT New isolated fusion accessory factor STRL33 - which mediates HIV  
XX PT infection of cells, used to develop products for the study, prevention  
XX PS and therapy of HIV-related disorders.  
XX PS Claim 65; Fig 4; 74pp; English.  
XX CC This sequence encodes the human STRL33 polypeptide that is expressed by  
XX CC the recombinant cell line of the invention. The STRL33 polypeptide is a  
XX CC fusion accessory factor that can function with CD4 to mediate fusion with  
XX CC cells bearing HIV-1 env from both laboratory-adapted TCM-tropic variants  
XX CC and from M-tropic variants. Agents that suppress STRL33 can be used in a  
XX CC method to treat HIV-related disorders. The products and methods can be  
XX CC used for the study, prevention and therapy of HIV-related disorders  
XX SQ Sequence 1918 BP; 493 A; 453 C; 444 G; 528 T; 0 U; 0 Other;  
Query Match 56.3%; Score 704.8; DB 2; Length 1918;  
Best Local Similarity 78.5%; Pred. No. 9.3e-209;  
Matches 870; Conservative 0; Mismatches 232; Indels 6; Gaps 2;  
QY 141 TTCACAAATTCACGTATACAGCC---AGGAGAACAAAGCTTCTCTAAAGTTCAAGGAG 197  
DB 64 TTCAGCATTTCAATGACAGCAGCAGGAGAGCATCAAGACTTCTGCAATTGAGCAAG 123  
QY 198 GTCTTTTGGCCTGTGTACCTGGTAGTGTGTTGCTGTTGAGCTGTAGGAACTCCCTG 257  
DB 124 GTCTTTCTGCCCTGCATGTACCTGGTGTGTTGCTGTTGCTGTTGTTGTTGTTGTTGTT 183  
QY 258 GTTCTGATATATACATTTTCTACCAAGCTGAGGACTCTGACAGATGTTTCTGCTG 317  
DB 184 GTCTGGTCAATCCATCTTCTACCATAGTTGAGAGCCTGAGGATGTTTCTGCTG 243  
QY 318 AACTTGCCCTGGCTGACCTGGTGTGTTGCTGTTACTCTGCTGCTTGGGCTATGAGGC 377  
DB 244 AACCTACCCCTGGCTGACCTGGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 303  
QY 378 ACCTATGAGTGGGTCTTGGCACAGTCACTGCAAACTCTTGGAGCATGTATACATG 437  
DB 304 ATCCATGAATGGGTGTTGGCCAGGTGATGTGAAGAGCCTACTGGCATCTACATTT 363  
QY 438 AACTTCTACGTGCTCCATGCTCACTCTCACTGATCACTGATGATGCTTTTCAATTGATG 497





Db 434 TGACTTGGGGAGGTCTATCTGCTGCTCATCTGGGTGATATCCCTGCTGGTTCCTTGC 493  
 QY 598 CACAGATCATCTATGGGCATGTTCAAGATATGTGAAGCTTATCTGTCAAGTACCAAGTG 657  
 Db 494 CCCAAATATCTATGGCAATGCTTTTAACTGGACAAGCTCATATGTGGTTATCATGACA 553  
 QY 658 AGGAGATATCCACTATGTTCTGTATACAGATGACTCTGGGTCTTCTCTGCCATTGC 717  
 Db 554 AGGAGATTTCCACTGTGGTCTTGGCCACCAGATGACACTGGGTCTTCTTGGCACTGC 613  
 QY 718 TCACATGATCTGTGCTACTCAGGCATTTATCAAGACTCTGCTTCATGCTCGAAACTTCC 777  
 Db 614 TCGCATGATGCTGCTATTCAGTCAATATCAAAACACTGCTTCATGCTGGAGGCTTCC 673  
 QY 778 AGAGCACAAATCTCTAAGATCATCTCTGTAGTGGCTGTGTTCTCTGCTGACCCAGA 837  
 Db 674 AGAAGCACAGATCTCTAAGATCATCTCTGTGATGGCTGTGTTCTCTGCTGACCCAGA 733  
 QY 838 CACCTTTCAACCTTGCCATGTTATCCAAAGTACAAAGTCAAGCTGGAGTACTATACCATTAACCA 897  
 Db 734 CACCTTTCAACCTGCTGAGCTCATCCGACAGACACACTGGAGTACTATGCCATGACCA 793  
 QY 998 GCTTTAGTATGCCATGCTATGAGTACAGAGGCTATAGCATCTTTCGGGCTTGCCTTAACC 957  
 Db 794 GCTTTACTACACCATCATAGTACAGAGGCCATCGCATACCTGAGGGCTGCTGCTTAACC 853  
 QY 958 CTGTACTTTATGCTTTGTTGGCTTAAAGTTCCGGAAGACGCTCGAACTTATCAAGG 1017  
 Db 854 CTGTGCTATGCTTTGCTGAGCTGAAGTTTGAAGAATCTTGAAGAATCTTGAAGA 913  
 QY 1018 ATATGCGCTGCTCTCTACCTGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTT 1077  
 Db 914 ACATTGCTGTCTCCCTTACCTTGGGCTCTCATCAATGGAATCTTCTGAGGACAAAT 973  
 QY 1078 CCAGAGCTTCTGCTGCCCAATGTAGAGACCAAGTATGTTCCAAATGTAG 1133  
 Db 974 CCAGAGCTTTTCTGCCCTCCCAATGTAGAGGCCACCAAGCATGTTCCAGTTATAG 1029

RESULT 8  
 ID AAS07616  
 XX AAS07616 standard; cDNA; 1029 BP.  
 AC AAS07616;

XX 23-OCT-2001 (first entry)  
 DE Pigtailed macaque cDNA encoding the SIV translocation agent, Bonzo.

XX Pigtailed macaque; Bonzo; simian immunodeficiency virus;  
 KW SIV translocating agent; CD4; BOB; Brother of Bonzo;  
 KW acquired immunodeficiency syndrome; AIDS; SIV; HIV; ss.

XX Macaca nemestrina.

PH Key Location/Qualifiers  
 FT CDS 1..1026  
 FT /\*tag= a  
 FT /product= "Bonzo"

XX US6251582-B1.

XX 26-JUN-2001.

XX 16-JUL-1998; 98US-00116498.

XX 17-JUL-1997; 97US-0052827P.

XX (UUNY) UNIV NEW YORK STATE.

XX Littman DR, Deng H, Unutmaz D, Kewalramani VN;

XX

DR WPI; 2001-424455/45.  
 DR P-PSDB; AAU04034.

XX Identifying an agent for treating AIDS comprises administering a  
 PT potential therapeutic agent to a cell expressing human CD4 and primate  
 PT Bonzo or BOB and contacting the cell with a virus pseudotyped with an HIV  
 PT envelope glycoprotein.

XX Example 1; Col 43-44; 41pp; English.

CC The sequence encodes a novel SIV (Simian immunodeficiency virus)  
 CC translocating agent, Bonzo. The invention relates to selecting an agent  
 CC that may be used in treating acquired immunodeficiency syndrome (AIDS)  
 CC comprises administering a potential therapeutic agent to a cell  
 CC expressing human CD4 and primate Bonzo or BOB (Brother of Bonzo),  
 CC contacting the cell with a virus pseudotyped with an HIV envelope  
 CC glycoprotein, and measuring the cell's ability to resist infection. The  
 CC method is useful for selecting or identifying an agent, which can be used  
 CC in treating AIDS

XX Sequence 1029 BP; 230 A; 278 C; 236 G; 285 T; 0 U; 0 Other;

Query Match 55.1%; Score 689.4; DB 5; Length 1029;  
 Best Local Similarity 79.8%; Pred. No. 4.2e-204;  
 Matches 811; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 118 ACTACGAGGAGATTCTGGCTCTTCAACATTCAGTATACAGCCAGGAGAACAAAC 177  
 Db 14 ATTACATGAAGACTATGGGCTCAACAGTTTCAATGACAGCCAGGAGGAGCATCAAG 73  
 QY 178 GCTTCTCTAAAGTTCAAGAGAGTCTTTTGGCCCTGTGTACCTGGTAGTGTGTTCTTTG 237  
 Db 74 ACTTCTGCTAGTTCAGGAAGTCTTTCTGCTGCTGATGTACCTGGTGTGTTGCTGTG 133  
 QY 238 GACTGCTAGGAACTCCCTGCTGCTGATTATATACATTTCTACCAGAACTGAGGACTC 297  
 Db 134 GCTTGGTGGGAACTCCCTGCTGCTGATATCCATCTTCTACCATAACTGCGAGGCC 193  
 QY 298 TGACAGATGTGTTTCTGCTGAACCTTGCCCTGGCTGACCTGGTGTGTTGTCTACTCTGC 357  
 Db 194 TGACGGAGCTGTTCTGCTGTAACCTACCCCTGGCTGACCTGGTGTGTTGTCTGACTCTGC 253  
 QY 358 CCTTTGGGCTATGACAGGACCTATGAGTGGGTCTTTGGCACAGTATGTGCAAAACTC 417  
 Db 254 CCTTCTGGGCTATGACAGGATCCATGAATGGATCTTTGGCCAGGTCTATGCAAGACCC 313  
 QY 418 TTGAGGAGCATGTATACAAATGAATCTTACGTGCTCCTGCTCACTCTCACCTGCATCACAG 477  
 Db 314 TACTGGGCTCTACACTATTAACTTTCTACATCCATGCTCATCTCCCTGCATCACTG 373  
 QY 478 TGGATCGTTTCATTGTAGTGTCCAGGCTACAGGCTTCAAGGCTTCAACCGGAGGTGAAGTGA 537  
 Db 374 TGGATCGTTTCATTGTAGTGTAAAGGCCCAAGGCTTCAACAGGAGGAGGAGGA 433  
 QY 538 AGATCTGGGCGCAAGTCAATTTGCTGCTCATTTGGGTGGTCTCCCTGTTGGTTTCTTTCG 597  
 Db 434 TGACTTGGGCAAGGTCTATCTGCTCATCTGGTGATATCCCTGCTGTTTCTTTCG 493  
 QY 598 CACAGATCATCTATGGCCATGTTTCAAGATATTGACAAGCTTATCTGTCAAGTACACAGTG 657  
 Db 494 CCCAAATATCTATGGCAATGTCTTTAACTTGACAAGCTCATATGTGGTTATCATGACA 553  
 QY 658 AGGAGATATCCATATGGTCTTGTATACAGATGACTCTGGGGTCTTCTCTGCCATTGC 717  
 Db 554 AGGAGATTTCCACTGTGGTCTTGGCCACCCAGATGACACTGGGGTCTTCTTCTGCCACTGC 613  
 QY 718 TCATATGATTCTGTGCTACTCAGGCATTTATCAAGACTTTCATGCTTCGAAACTTCC 777  
 Db 614 TCGCATGATTGCTGCTATTTCAGTCAATCAAAACACTGCTTTCATGCTGGAGGCTTCC 673  
 QY 778 AGAAGCAAAATCTCTAAGATCATCTTCTTGTAGTGGCTGTGTTCTCTGCTGACCCAGA 837  
 Db 674 AGAAGCACAGATCTCTAAGATCATCTTCTTGTGTGGTGTGTTCTCTGCTGACCCAGA 733

QY 838 CACCTTCACCTGCGATGTTAATCCAAAGTACAAAGCTGGAGTACTATACCATAACCA 897  
 Db 734 CACCTTCACCTGCGAGCTCATCCGAGCAGACACACTGGGAGTACTATGCCATGACCA 793  
 QY 898 GCTTTAAGTATGCGATCGTAGTACAGAGGCTATAGCATACTTTCGGGCTTGCCTTAACC 957  
 Db 794 GCTTTCACATACATATAGTACAGAGGCTATGCTATACCTGAGGGCTGCTTAACC 853  
 QY 958 CTCTACTTATGCTTGTGCTTAAAGTTCCGGAAGACCTCTGGAACCTTATGAAGG 1017  
 Db 854 CTGTGCTCTATGCTTGTGCTTAAAGTTCCGGAAGACCTCTGGAACCTTATGAAGG 913  
 QY 1018 ATATCGGCTGCTCTCTACCTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACATTT 1077  
 Db 914 ACATGGCTGTCTCTCTACCTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACATTT 973  
 QY 1078 CCAAGACTTGTCTGCTCTCCCAATGATGAGACACCAAGTATGTTCCAAATGTAG 1133  
 Db 974 CCAAGACTTGTCTGCTCTCCCAATGATGAGACACCAAGTATGTTCCAAATGTAG 1029

## RESULT 9

ABN84641  
 ID ABN84641 standard; cDNA; 1029 BP.

XX AC ABN84641;

XX XX 28-OCT-2002 (first entry)

XX XX Pig-tailed macaque G-protein coupled receptor Bonzo cDNA.

XX KW Bonzo; G-protein coupled receptor; receptor; STRL33; TYMSR; retrovirus;  
 XX KW translation promoting agent; anti-HIV; antiviral; virucide; AIDS;  
 XX KW antiinflammatory; immunostimulant; transgenic animal; pig-tailed macaque;  
 XX KW gene; ss.

XX OS Macaca nemestrina.

XX FH Key Location/Qualifiers

XX CDS 1..1029

XX FT /\*tag= a

XX FT /product= "Bonzo"

XX PN US2002076694-A1.

XX XX 20-JUN-2002.

XX XX 09-MAY-2001; 2001US-00852156.

XX XX 17-JUL-1997; 97US-0052827P.

XX XX 16-JUL-1998; 98US-00116498.

XX XX (LITT/) LITTMAN D R.

XX XX (DENG/) DENG H.

XX XX (UNUT/) UNUTMAZ D.

XX XX (KEMA/) KEWALRAMANI V N.

XX XX Littman DR, Deng H, Unutmaz D, Kewalramani VN;

XX XX WPI; 2002-598719/64.

XX XX P-PSDB; ABN79745.

XX XX New isolated nucleic acid encoding the translocation promoting agent  
 XX XX Bonzo, useful e.g. in screening for agents that inhibit infection by  
 XX XX human immune deficiency virus.

XX XX Disclosure; Page 31; 65pp; English.

XX XX The present sequence is the coding sequence for pig-tailed macaque Bonzo  
 XX XX (also named STRL33 or TYMSR). Human Bonzo (see ABN79743) has been  
 XX XX identified as a novel G-protein coupled receptor involved in retroviral  
 XX XX entry into cells, acting in conjunction with CD4 as a translocation

CC promoting agent for SIV and also for certain strains of HIV-2 and M-  
 CC tropic HIV-1. Bonzo and the closely related BOB (brother of Bonzo) are  
 CC related to the chemokine receptor family and are expressed in lymphoid  
 CC tissues. Use of these new receptors following experimental infection of  
 CC non-human primates with SIV strains may provide important insight into  
 CC viral transmission and mechanisms of SIV- and HIV-induced AIDS. The  
 CC invention provides Bonzo and BOB amino acid and DNA sequences, mammalian  
 CC cells transfected with Bonzo and/or BOB and human CD4, and antibodies to  
 CC Bonzo. A method of identifying other translocation promoting agents is  
 CC also disclosed. Cells (or transgenic animals) that express Bonzo or BOB  
 CC and CD4 are used to: (a) identify therapeutic agents for treating AIDS  
 CC (acquired immune deficiency syndrome), by increasing the resistance of  
 CC cells to infection; (b) in filters to remove viruses with HIV envelope  
 CC glycoproteins from biological fluids; (c) to identify ligands for Bonzo  
 CC or BOB; (d) to identify compounds that increase the immune response to a  
 CC specific pathogen or vaccine by increasing Bonzo/STRL33 expression; and  
 CC (e) to identify antiinflammatory agents that reduce Bonzo/STRL33  
 CC expression and thus inhibit recruitment of memory cells

XX SQ Sequence 1029 BP; 230 A; 278 C; 236 G; 285 T; 0 U; 0 Other;

Query Match 55.1%; Score 689.4; DB 6; Length 1029;

Best Local Similarity 79.8%; Pred. No. 4.2e-204;

Matches 8.1; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 118 ACTACGAGGAGATTTCTGGCTCTTCAACAATTCAGTGTATACAGCAGAGAACAAAC 177

Db 14 ATTACCATGAAGACTATGGGCTCAACAGTTTCAATACAGCAGCAGGAGGAGCATCAAG 73

QY 178 GCTTCTCTAAAGTTCAAGGAGGTCCTTTTGGCTGTGTACTGTGTACTGTGTCTTTGCTTTG 237

Db 74 ACTTCTCTCAGTTTCAAGGAGGTCCTTTTGGCTGTGTACTGTGTACTGTGTCTTTGCTTTG 133

QY 238 GACTGTCTAGAACTCCCTGGTCTGTGATATATACATTTTACAGAACTGAGGACTC 297

Db 134 GCTGTGTGGGAACTCCCTGGTGTGTGTATCCATCTTACCATAAAGTGTGAGAGCC 193

QY 298 TGACAGATGTGTCTGTGAACTTGCCTGGCTGACCTGGTGTCTGTCTGTACTCTGC 357

Db 194 TGACGAGCTGTCTGTGTGAACTTGCCTGGCTGACCTGGTGTCTGTCTGTACTCTGC 253

QY 358 CTTTGTGGGCTATGACGAGCCTATGAGTGGTCTTTGGCAGAGTCAATGTCGAAAATC 417

Db 254 CTTTGTGGGCTATGACGAGCCTATGAGTGGTCTTTGGCAGAGTCAATGTCGAAAATC 313

QY 418 TTGAGGAGATGTATACAAATGAACTTCTACGTGTCCATGCTCACTCCTCCTGCATCAG 477

Db 314 TACTGGGCTGTACACTATTAACCTTACACATCCATGCTCATCTCCTGCATCAGT 373

QY 478 TGGATCGTTTCAATGTAGTGGTCAAGGCTACAAAGGCTTCAACCGGAGGCTAAGTGA 537

Db 374 TGGATCGTTTCAATGTAGTGGTCAAGGCTTCAAGGCTTCAACCGGAGGCTAAGTGA 433

QY 538 AGATCTGGGGCAAGTCAATTTGCTCTCATTGGGTGGTCTCCCTGTGTTTCTTTGC 597

Db 434 TGACTTGGGGCAAGTCAATTTGCTCTCATTGGGTGGTCTCCCTGTGTTTCTTTGC 493

QY 598 CACAGATCATCTATGGCCATGTTCAAGATATTGACAAGCTTATCTGTCACTACACAGTG 657

Db 494 CCAAAATTTATCTATGGCAATGTTTAACTCTGGAACAAGCTCATATGTGTTTATCATGACA 553

QY 658 AGGAGATATCCACTATGCTTCTGTATACAGATGACTCTGGGGTCTCTCTGTCGCTTGC 717

Db 554 AGGAGATTTCCACTGCTGTTCTTGGCCACCCAGATGACACTGGGGTCTCTCTGTCGCTTGC 613

QY 718 TCATATGATTTCTGTGCTACTCAGGCATTTATCAAGACCTTGTCTCATGCTCGAACTTCC 777

Db 614 TCGCCATGATTTCTGTGCTACTCAGGCATTTATCAAAACACTGCTTCTCATGCTCGAACTTCC 673

QY 778 AGAAGCACAATCTCTAAAGATCATCTTCTGTAGTGGCTGTCTTCTGTGTGACCCAGA 837

Db 674 AGAAGCACAATCTCTAAAGATCATCTTCTGTAGTGGCTGTCTTCTGTGTGACCCAGA 733

QY 838 CACCCCTCAACCTTCCCATGTTATATCCAAAGTACAGCTGGAGTACTATACCATACCA 897  
 DB |||||  
 DB 734 CACCCCTCAACCTTCCCATGTTATATCCAAAGTACAGCTGGAGTACTATACCATACCA 793  
 QY 898 GCTTTAAGTATGCCATCGTACGTAGTACAGAGGCTATAGCATCTTTCGGGCTTGCTTAACC 957  
 DB |||||  
 DB 794 GCTTTCACTACACCATCATAGTACAGAGGCTATAGCATCTTTCGGGCTTGCTTAACC 853  
 QY 958 CTGTACTTTATGCTTTGTTGGCTTAAAGTTCGGAAGAACGTCGTGAAACTTATGAAGG 1017  
 DB |||||  
 DB 854 CTGTGCTCTATGCTTTGTCAGCTCGAAGTTTCGAAAGAACTTCTGGAAGTGTGAAGG 913  
 QY 1018 ATATCGGCTGCTCTCTCACTCGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTT 1077  
 DB |||||  
 DB 914 ACATGCTGCTCTCTCTCACTCGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTT 973  
 QY 1078 CCAAGACTTGTCTGCTCCACAAATGTAGACACCAAGTATGTTCCAAATTGTAG 1133  
 DB |||||  
 DB 974 CCAAGACTTGTCTGCTCCACAAATGTAGACACCAAGTATGTTCCAAATTGTAG 1029

## RESULT 10

AXX07290  
 ID AAX07290 standard; cDNA; 1037 BP.

XX AC AAX07290;  
 XX AC

XX DT 21-MAY-1999 (first entry)

XX DE African green monkey HIV/SIV receptor protein Bonzo cDNA.

XX KW SIV receptor; HIV receptor; Bonzo; translocation promoting agent;  
 KW G protein coupled receptor; chemokine receptor; AIDS; infection;  
 KW virus transmission; African green monkey; ds.

XX OS Cercopithecus aethiops.  
 XX OS

XX PN W09903888-A1.  
 XX PN

XX PD 28-JAN-1999.  
 XX PD

XX PF 17-JUL-1998; 98WO-US014857.  
 XX PF

XX PR 17-JUL-1997; 97US-00896155.  
 XX PR

XX PA (UJNY) UNIV NEW YORK STATE.  
 XX PA

XX PI Littman DR, Deng H, Unutmaz D, Kewalramani VN;  
 XX PI

XX DR WPI; 1999-132164/11.  
 XX DR

XX DR P-PSDB; AAW97784.

PT New nucleic acid encoding the human translocation promoter Bonzo - used  
 PT to screen for potential agents for treating acquired immune deficiency  
 PT syndrome.

PS Disclosure; Page 79-80; 97pp; English.  
 PS

XX This cDNA clone codes for African green monkey Bonzo (see AAW97784), a  
 CC novel HIV/SIV translocation promoting agent that acts in conjunction with  
 CC CD4 to serve as a receptor for the entry into a cell of a virus having a  
 CC specific viral envelope glycoprotein. Bonzo is a member of the 7-  
 CC transmembrane G-protein coupled receptor family, is closely related to  
 CC the chemokine receptor family, and is expressed in lymphoid tissues. The  
 CC invention provides the amino acid sequences (see AAW97783-88) and DNA  
 CC sequences (see AAX07289-94) of human, African green monkey and pig-tailed  
 CC macaque Bonzo and Bob (brother of Bonzo) translocation promoting agents.  
 CC These novel receptors were identified using an expression cloning  
 CC strategy. They were found to be used by particular strains of HIV-2 and M-  
 CC tropic HIV-1. The invention also includes: mammalian cells transfected  
 CC with Bonzo and/or Bob and human CD4, which can be used to screen  
 CC potential therapeutic agents and identify ligands; antibodies to Bonzo,  
 CC which can be used therapeutically, e.g. as antagonists or to target

CC toxins or radioisotopes to HIV-permissive cells; transgenic animals; and  
 CC antisense and ribozyme molecules, which may also be used therapeutically,  
 CC particularly expressed from a gene therapy vector

XX SQ Sequence 1037 BP; 232 A; 276 C; 235 G; 294 T; 0 U; 0 Other;

Query Match 54.8%; Score 685; DB 2; Length 1037;  
 Best Local Similarity 81.1%; Pred. No. 9.9e-203;  
 Matches 809; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

QY 140 CTTCAACAAATTCAGTGATACAGCCAGGAGAACAAAC--GCTTCCTAAAGTTCAAGGA 196  
 DB |||||

QY 197 GGTCTTTTGGCCCTGTGTACTACCTGGTAGTGTGTCTTTGGACTGTCTAGGAAATCCCT 256  
 DB |||||

QY 257 GGTTCGTGATTATATACATTTTCTACCAAGAGTGAAGACTCTGACAGATGTTTCTGCT 316  
 DB |||||

QY 317 GAACTTGCCCTGGCTGACCTGGTGTGTGTCTGTACTCTGCTCTTTGGGCTATGCAGG 376  
 DB |||||

QY 377 CACCTATGAGTGGGTCTTTGGCACAGTCATGTGCAAAACTCTTCGAGGCATGTATCAAT 436  
 DB |||||

QY 437 GAACTTCTACGTGTCCATGCTCACTCTCACCTGCATCACAGTGGATCGTTTCAATTGAGT 496  
 DB |||||

QY 497 GGTCCAGCTACCAAGGCTTCAACGGGAGGCTTAAGTGAAGATCTGGGCGCAAGTCAT 556  
 DB |||||

QY 557 TTGCTTGTCTCATTTGGGTGCTCCCTGTTGGTTCCTTTCGCCACAGATCATCTATGCCA 616  
 DB |||||

QY 617 TGTTCAAGATATTGACAAAGCTTATCTGTCTAGTACCACAGTGAAGAGATATCCACTATGT 676  
 DB |||||

QY 677 TCCTTGTATACAGATGACTCTGGGTTCTTCTGCTGCCATTCCTCAGTATGATTCGTGCTA 736  
 DB |||||

QY 737 CTCAGGCATTATCAAGACCTTGTCTTCATGCTCGAAACTTCCAGAAGCACAAATCTCTAA 796  
 DB |||||

QY 797 GATCATCTCTCTGTAGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 856  
 DB |||||

QY 857 GTTAATCCAAAGTACAGCTGGGAGTACTATACATAACCCAGCTTTAAGTATGCCATCGT 916  
 DB |||||

QY 916 AGTGACAGAGGCTATAGCATATCTTTCGGGCTTGGCTTAAACCCCTGTCTCTATGCCCTTGT 976  
 DB |||||

QY 977 TGGCTTAAAGTTCGGAAGAACGCTCTGGAACCTTATGAAGGATATCGGCTGCTCTCA 1036  
 DB |||||

QY 1037 CCTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTCAGAGCTTGTGTGCTTC 1096  
 DB |||||

QY 933 CCTTGGGCTCTCATCATCAATGGAATCTTCTGAGGACAAATTCAGAGCTTTTCTGCTTC 992  
 DB |||||



antiinflammatory; immunostimulant; transgenic animal;  
African green monkey; gene; ss.

*Cercopithecus aethiops*.

Key	Location/Qualifiers
CDS	1. .1029
	/*tag= a
	/product= "BenzO"

US2002076694-A1.

20-JUN-2002.

09-MAY-2001: 2001US-00852156.

17-III.-1997: 97US-0052827P

17-OCT-1997; 97US-0032827F;  
16-JUL-1998; 98US-00116498;

(LITT/) LITTMAN D R.

(DENG//) DENG H.

(UNUT /) UNUTMAZ D.

(KEWA/) KEWALRAMANI V N.

Littman DR, Deng H, Unutmaz D, Kewalramani VN;

WPI; 2002-598719/64.

P-PSDB; ABB79744.

New isolated nucleic acid encoding the translocation promoting agent Bonzo, useful e.g. in screening for agents that inhibit infection by human immune deficiency virus.

Disclosure; Page 29; 65pp; English.

The present sequence is the coding sequence for African green monkey Bonzo (also named STRL33 or TYMSTR). Human Bonzo (see AB379743) has been identified as a novel G-protein coupled receptor involved in retroviral entry into cells, acting in conjunction with CD4 as a translocation promoting agent for HIV and also for certain strains of HIV-2 and M-tropic HIV-1. Bonzo and the closely related BOB (brother of Bonzo) are related to the chemokine receptor family and are expressed in lymphoid tissues. Use of these new receptors following experimental infection of non-human primates with HIV strains may provide important insight into viral transmission and mechanisms of HIV- and HIV-induced AIDS. The invention provides Bonzo and BOB amino acid and DNA sequences, mammalian cells transfected with Bonzo and/or BOB and human CD4, and antibodies to Bonzo. A method of identifying other translocation promoting agents is also disclosed. Cells (or transgenic animals) that express Bonzo or BOB and CD4 are used to: (a) identify therapeutic agents for treating AIDS (acquired immune deficiency syndrome), by increasing the resistance of cells to infection; (b) in filters to remove viruses with HIV envelope glycoproteins from biological fluids; (c) to identify ligands for Bonzo or BOB; (d) to identify compounds that increase the immune response to a specific pathogen or vaccine by increasing Bonzo/STRL33 expression; and (e) to identify antiinflammatory agents that reduce Bonzo/STRL33 expression and thus inhibit recruitment of memory cells.

Sequence 1037 BP: 232 A: 276 C: 235 G: 294 T: 0 U: 0 Other:

54.8%: Score 685: DB 6: Length 1037:

54.8%; score 685; DB 6; Length 1037  
 at Local similarity 81.1%; Pred. No. 9.9e-203;

Sequence	Local Similarity	Pred. No.	Gaps
Conservative	81.1%	9.9e-203	1
Mismatches	809	0	3
Indels	185	0	3

140 CTTCAACAATTCCAGTGATAACAGCCAGGAGAACAAAC--GCTTCCTAAAGTCAAGGA 196

[illegible]

197 GGTCTTTTGCCTGTCGTACCTGGTAAGTCTTCTCTTCTAACAATAATCCCT  
256

0  
1  
2  
3  
4  
5  
6  
7  
8  
9  
A  
B  
C  
D  
E  
F  
G  
H  
I  
J  
K  
L  
M  
N  
O  
P  
Q  
R  
S  
T  
U  
V  
W  
X  
Y  
Z  
[  
\  
]  
^  
\_  
`  
a  
b  
c  
d  
e  
f  
g  
h  
i  
j  
k  
l  
m  
n  
o  
p  
q  
r  
s  
t  
u  
v  
w  
x  
y  
z  
{  
|  
}  
~  
\_

[illegible]

Db	153	GGTGTGGTGCATATCCATCTTTCTACCATAAAGCTACAGAGCCTGACGGACGTGTTCTCTGGT	213
Qy	317	GAACTTGGCCCTGGCTGACCTGGTGTGTGTCTGTACTCTGCGCTTTTGGGCGCTATGCAGG	376
Db	213	GAACTTACCCCTGGCTGACCTGGTGTGTGTCTGTACTCTGCGCTTTTGGGCGCTATGCAGG	272
Qy	377	CACCTATGAGTGGGTCTTTTGGCAGCAGTCATGTGCAAAACTCTTCGAGGCAATGTATACAAT	436
Db	273	CATCCATGAATGGATCTTTTGGCCAGGTCATGTGCAAGACCCTACTGGGTATCTACACTAT	332
Qy	437	GAACTTCTACGTGTCCATGTCTCACTCTCACTCTGCATCACAGTGCATCGTTTTCATTTGAGT	496
Db	333	TAACTTCTACACATCTATGTCTCATCCCTCACTGCATCACTGTGCATCGTTTTCATTTGAGT	392
Qy	497	GGTCCAGGCTACCAAGGCGCTTCAACCGCAGGCTTAAGTGAAGATCTCGGGCCCAAGTCAT	556
Db	393	GGTTAAGGCCACCAAGGCCATATACCAAGCAAGCCCAAGAAGATGACTTGGGGCAAGTTCAT	452
Qy	557	TTGCTTTCGTCTAATTTGGGTGGTCTCCCTGTGTGGTTCTTTTGGCCACAGATCACTATAGGCCA	616
Db	453	CTGCTTTCGTCTAATTTGGGTGGTATCCCTGTGTGGTTCTTTCGCCAAATATCTATGCCAA	512
Qy	617	TGTTCAAGATATTACACAAGCTTATCTCTCAGTACCAAGTACCAAGTACGAGAGATATCCACTATGGT	676
Db	513	TGCTTTTAATCTGGACAAGCTCATATGTGTGTATCATGTATGAGAGATTTCCACTGTGGT	572
Qy	677	TCTTGTATTACAGATGACTCTGGGGTTCTTTCTGGCCATTCCTCACTATGATTTCTGTGCTA	736
Db	573	TCTTGGCACCAGATGACACTGGGGTTCTTTCTTGGCCACTGCTGCCATGATTTGTCTGCTA	632
Qy	737	CTCAGGATTTATCAAGACCTTGTCTCATCTCGAACTTCGAAATTCAGAGACACAAATCTCTAAA	796
Db	633	TTTCACTAATCAAAACACTGTCTTCACTGTGGAGGCTCCAGAGCACAGATCTCTAAA	692
Qy	797	GATCATCTTCTTTGTAGTGGCTGTGTTCTCTGTGACCCAGACACCCCTTCAACCTTGGCAT	856
Db	693	GATCATCTTCTTTGTGATGCTGTGTTCTCTGTGACCCAGACACCCCTTCAACCTCGTAA	752
Qy	857	GTTAATCCAAAGTACAAGCTGGGAGTACTATACATAACCAGCTTTAAGTATGCCATCGT	916
Db	753	GCTCATCCGAGCACACACTGGGAGTACTATGCCATGACCAGCTTTCACATACACCATCAT	812
Qy	917	AGTGACAGAGCTATAGCATACTTTCGGGCTTGCCTTAACCCGTACTTTATGCTTTGT	976
Db	813	AGTGACAGAGCCATCGCATACCTGAGGGCTTGCCTTAACCCGTGCTCTATGCCTTGT	872
Qy	977	TGGCTTAAAGTTCGGAAGAACCTCTGGAAACTTATGAAGGATATCGGTGCTCTCTCA	1036
Db	873	CAGCCTGAAGTTTCGAAAGAACTTCTGGAACTTGTGAAGACATTTGGCTGTCTCCCTTA	932
Qy	1037	CTTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTCCAAGACTTGTTCCTGCTC	1096
Db	933	CCTTGGGCTCTCATCAATGGAACTTCTCTGAGGACATTTCCAGACATTTTCTGCCCTC	992
Qy	1097	CCACAATGTAGACACCAGTATGTTTCCAAATTTGTAG	1133
Db	993	CCACAATGTGGAGGCCACCAGCATGTTTCCAGTTATAG	1029
RESULT 13			
ABS73339			
ID ABS73339 standard; cDNA; 1029 BP.			
XX	ABS73339;		
XX	04-DEC-2002 (first entry)		
DT	cDNA encoding human GPCR STLR33.		
XX	Human; transmembrane receptor; G-protein coupled receptor; GPCR; allergy;		
KW	hypertension; reflux disease; depression; migraine; schizophrenia; ulcer;		
KW	psychotic disorder; asthma; bronchospasm; anaesthesia;		

KW myocardial infarction; MI; stroke; glaucoma; anxiety;  
 KW prostatic hyperplasia; epilepsy; prostate cancer; rhinitis; angina;  
 KW prostatic hypertrophy; gene; ss.  
 XX Homo sapiens.  
 XX WO300268600-A2.  
 XX PD 06-SEP-2002.  
 XX PF 26-FEB-2002; 2002WO-US005625.  
 XX PR 26-FEB-2001; 2001US-0271913P.  
 XX PA (AREN-) ARENA PHARM INC.  
 PI Liaw CW, Chalmers DT, Behan DP, Maciejewski-Lenior D, Leonard JN;  
 PI Lin I, Ortuno D;  
 XX WPI; 2002-706980/76.  
 DR P-PSDB; ABG95153.  
 XX New human G-protein coupled receptor (GPCR), useful for screening agonist  
 PT or inverse agonist compounds for treating diseases associated with GPCR.  
 XX Claim 7; Page 107-108; 201pp; English.  
 XX The present invention relates to transmembrane receptors, particularly  
 CC endogenous human G-protein coupled receptors (GPCRs), mutant (non-  
 CC endogenous) versions of the GPCRs, and the polynucleotide sequences  
 CC encoding them. The GPCRs are useful for screening agonist or inverse  
 CC agonist compounds for treating diseases associated with GPCR. Diseases  
 CC that can be treated with such compounds include allergies, hypertension,  
 CC reflux disease, depression, migraine, schizophrenia, ulcers, psychotic  
 CC disorders, asthma, bronchospasm, anaesthesia, myocardial infarction (MI),  
 CC stroke, glaucoma, prostatic hyperplasia, epilepsy, prostate cancer,  
 CC anxiety, prostatic hypertrophy, rhinitis, and angina. The present  
 CC sequence encodes an endogenous human GPCR  
 XX Sequence 1029 BP; 229 A; 276 C; 237 G; 287 T; 0 U; 0 Other;

Query Match 54.4%; Score 680.8; DB 6; Length 1029;  
 Best Local Similarity 81.3%; Pred. No. 2e-201;  
 Matches 787; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 166 AGGAGAACAAACGCTTCTAAAGTTCAAGGAGGCTTTTTCCTGCTGTGTACCTGTAG 225  
 DB 62 AGGACATCAAGCCTTCTCGAGTTTCAGCAAGGCTTTTCTGCTGCTGCTGCTGCTG 121  
 QY 226 TGTGTGCTTTGGACTGTAGGAACTCCCTGTTCTGATTATATACATTTTCTACACAGA 285  
 DB 122 TGTGTGCTGTGCTGTGGGAACTCTCTGGTCTGTTGCTATATCACTTCTACCATTA 181  
 QY 286 AGCTGAGGACTGTGACAGATGTTTCTGCTGAATTCCTGCTGCTGCTGCTGCTGCTGTTG 345  
 DB 182 AGTTGACAGGCTGACGAGATGTTTCTGCTGTAACCTACCTGCTGCTGCTGCTGCTGTTG 241  
 QY 346 TCTGTACTCTGCTCTTTGGGCTATGACGACCTATGAGTGGTCTTGTGSCACAGTCA 405  
 DB 242 TCTGACATCTGCTCTTCTGGGCTATGACGAGCATCCATGATGGTGTGTGGCCAGGTCA 301  
 QY 406 TGTGCAAACTCTTCGAGGCATGTATACAACTTCTACGTGTCCATGTCTCACTCTCA 465  
 DB 302 TGTGCAAAAGCCTACTGGGCATCTACACTATTAACTTCTACAGCTCCATGTCTCACTCA 361  
 QY 466 CTGTGATCAAGTGTGCTTTTCAATGTAGTGGTCCAGCTACCAAGGCTTCAACGGC 525  
 DB 362 CTGTGATCACTGTGATGCTTTTCAATGTAGTGGTAAAGGCCACCAAGGCTTCAACAGC 421  
 QY 526 AGCTTAAGTGGAGATCTGGGGCCCAAGTCAATTTGCTCATTTGGGTGGTCTCCCTGT 585  
 DB 422 AAGCCAAAGAGATGACCTGGGGCAAGGTCACCAAGCTTCTCATCTGGGTGATACCTGTC 481

QY 586 TGGTTTCTTTGGCCACAGATCACTATATGGCCATGTTCAAGATATTTGACAAGCTTATCTGTC 645  
 DB 482 TGGTTTCTTTGGCCCAAAATATCTATATGGCAATGCTTTTAAATCTCGACAAGCTCATATGTG 541  
 QY 646 AGTACCACAGTGGAGAGATATCCACTATGTTCTTGTATATACAGATGACTCTGGGGTTCCT 705  
 DB 542 GTTACCATGACGAGGCAATTTCCACTGTGTGTTCTTCCACCCAGATGACACTGGGGTTCCT 601  
 QY 706 TCCTGCAATTGCTCACTATGATTTCTGTCTACTCAGGCATTATCAAGACTTGTCTTCATG 765  
 DB 602 TCTTGCCACTGCTCACCATGATTTGCTGCTATTCACTATAATCAAAACACTGCTTCATG 661  
 QY 766 CTCGAAACTTCCAGAAGCAAAATCTCTAAAGATCATCTTCTCTGTAGTGGCTGTGTTCC 825  
 DB 662 CTGGAGGCTTCCAGAAGCAGATCTCTAAAGATCATCTTCTGTGGTGTGGCTGTGTTCC 721  
 QY 826 TGTGACCCACAGACACCCCTTCAACCTTGGCCATGTTATCCAAAGTACAAGCTGGGAGTACT 885  
 DB 722 TGCTGACCCAGATGCCCTTCAACCTCATGAGTTTCATCCGACGACACACTGGGAATACT 781  
 QY 886 ATACCATAACCAAGCTTTAAGTATGCCATCGTAGTGAAGAGGCTATAGCATATTTTCGGG 945  
 DB 782 ATGCCATGACCAAGCTTTTCACTACACCATCATGCTGACAGAGGCCATCGCATACCTGAGG 841  
 QY 946 CTGTCCTTAACCCCTGACTTTTATGCTTGTGGCTTAAAGTTCCGGAAGACGCTCTGA 1005  
 DB 842 CTGTCCTTAACCCCTGCTGCTATGCTTGTGAGCTGAGTTTCGAAGAACTTCTGA 901  
 QY 1006 AACTTATGAAGGATATCGGCTGCTCTCTCACTGGGAGTCTCAAGTCAATGAAGTCTT 1065  
 DB 902 AACTTGTGAAGGACATGTTGTTGCTCCCTTACCTGGGGTCTCACATCAATGGAATCTT 961  
 QY 1066 CTGAGCAGATTCGAAGACTTGTCTGCTGCCACATGTAGAGACCACGATGCTTCTC 1125  
 DB 962 CTGAGCAGATTCGAAGACTTGTCTGCTGCCACATGTAGAGACCACGATGCTTCTC 1021  
 QY 1126 AATTGTAG 1133  
 DB 1022 AGTTATAG 1029

RESULT 14  
 AAF85364  
 ID AAF85364 standard; cDNA; 1029 BP.  
 XX AAF85364;  
 XX 23-JUL-2001 (first entry)  
 DE Nucleotide sequence of a human Bonzo polypeptide.  
 KW Bonzo; CXCR chemokine receptor; inflammatory disease; cancer; infection;  
 KW SXCkine; spleen extracted chemokine; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 CDS 1..1029  
 FT /tag= a  
 FT /product= "Bonzo"  
 XX WO200137872-A1.  
 XX 31-MAY-2001.  
 XX 22-NOV-2000; 2000WO-US032206.  
 XX 24-NOV-1999; 99US-00449437.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX Briskin MJ, Murphy KE, Wilbanks AM, Wu L;

DR WPI; 2001-343947/36.  
 DR P-PSDB; AAB68422.  
 PT Identifying agents (especially antibodies) which bind to the CXCR  
 PT chemokine receptor Bonzo, and which may be used to treat e.g. cancers and  
 PT inflammation.  
 XX  
 PS Example 1; Fig 1; 191pp; English.  
 XX  
 CC The present sequence encodes a human Bonzo polypeptide. Bonzo is a CXCR  
 CC chemokine receptor. The specification describes a method for identifying  
 CC agents (especially antibodies) which bind to Bonzo and inhibit the  
 CC binding of a ligand (especially sexkine (spleen extracted chemokine))  
 CC and the agents per se. The agents identified may be used for the  
 CC treatment of a disorder/disease related to aberrant Bonzo expression and  
 CC activity, such as inflammatory disease, cancers and/or infections (e.g.  
 CC viral, bacterial and fungal infections)  
 XX  
 SQ Sequence 1029 BP; 229 A; 275 C; 238 G; 287 T; 0 U; 0 Other;  
 Query Match 54.3%; Score 679.2; DB 4; Length 1029;  
 Best Local Similarity 81.2%; Pred. No. 6.4e-201;  
 Matches 786; Conservative 0; Mismatches 182; Indels 0; Gaps 0;  
 QY 166 AGGAGAACAAACGGCTTCTTAAGTTCAAGGAGGTCCTTTTGGCCCTGTGTGCTACCTGGTAG 225  
 DB 62 AGGAGCATCAAGACTTCTCTCAGTTCAAGGAGGTCCTTTTGGCCCTGTGTGCTACCTGGTAG 121  
 QY 226 TGTGTTGCTTGGACTGTAGGAACCTCCCTGGTCTGATATATATATATATATATATATATATAT 285  
 DB 122 TGTGTTGCTTGGACTGTAGGAACCTCCCTGGTCTGATATATATATATATATATATATATATAT 181  
 QY 286 AGCTGAGGACCTGACAGATGTGTTCTGCTGAACCTGGCCCTGGTGTGCTGCTGCTGCTGCTG 345  
 DB 182 AGTTGACAGGCTGACGAGATGTGTTCTGCTGAACCTGGCCCTGGTGTGCTGCTGCTGCTGCTG 241  
 QY 346 TCTGTACTCTGCTTGGCCCTATGACGACCTATGAGTGGTCTTTGGCACAGTCA 405  
 DB 242 TCTGCACTCTGCTTGGCCCTATGACGACCTATGAGTGGTCTTTGGCACAGTCA 301  
 QY 406 TGTCAAAACCTTTCGAGGATGTATCAATGAATCTTACGTTCTCATCTCACTCTCA 465  
 DB 302 TGTCAAGAGCTTACTGGGATCTACACTATTAATCTTACGTTCTCATCTCACTCTCA 361  
 QY 466 CTGCACTACAGTGGATCGTTTCAATGAGTGGTCCAGGCTACCAAGGCTTCAACCGGC 525  
 DB 362 CTGCACTACAGTGGATCGTTTCAATGAGTGGTCCAGGCTACCAAGGCTTCAACCGGC 421  
 QY 526 AGCTAAGTGGAGATCTGGGCAAGTCAATTTGCTGCTCATTTGGTGGTCTCCCTGT 585  
 DB 422 AAGCAAGAGATGACCTGGGCAAGTCAATTTGCTGCTCATTTGGTGGTCTCCCTGT 481  
 QY 586 TGGTTTCTTGGCCACAGATCATCTATGGCCATGTTCAGATATTTGACAGCTTATCTGTC 645  
 DB 482 TGGTTTCTTGGCCACAAATATCATATGCAATGTCTTAACTCGAAGCTCATATG 541  
 QY 646 AGTACACAGTGGAGATATCACTATGTTTCTTGTATACAGATGACTCTGGGTTCT 705  
 DB 542 GTTACCATGAGGCAATTTCCACTGTGTTCTTGGCCACCCAGATGACACTGGGTTCT 601  
 QY 706 TCTGCACTTGTCTCATATGTTCTGTTACTCAGGCAATTTCAAGCTTCTGTTCTG 765  
 DB 602 TCTTGGCACTGTCTCACCATGTTTGTCTTATTCAGTCAATATCAAAACACTGTTCTG 661  
 QY 766 CTCGAAATCTCCAGAACGACAAATCTCTAAAGATCATCTTCTTGTAGTGGTGTGTTCC 825  
 DB 662 CTGAGGCTTCCAGAGACAGATCTCTAAAGATCATCTTCTTGGTGGTGTGTTCC 721  
 QY 826 TGTGACCCAGACACCTTCAACCTTCCCATGTTTATCCAAAGTACAAAGTGGGAGTACT 885  
 DB 722 TGTGACCCAGATGCTTCAACCTGATGAGTTCTCCGACACACACTGGGATATCT 781  
 QY 886 ATACCATACACAGCTTTAAGTATGCCATCGTAGTACAGAGGCTATAGCATACTTTCCGG 945

DB 782 ATGCCATGACAGCTTTTCACTACCATCATGTGTACAGAGCCATCCATCTAGGG 841  
 QY 946 CTGCTTTAAACCTCTACTTTTATGCTTTTAAAGTTCCGGAAGAACCTGTGA 1005  
 DB 842 CTGCTTTAAACCTCTGCTCTATGCTTTGTGAGCTTCTGAAAGAACTTCTGA 901  
 QY 1006 AACTTATGAGGATATCGGCTGCTCTCTACCTGGGAGTCTCAAGTCAATGGAATCTT 1065  
 DB 902 AACTTGTGAAGGACATGTTGGCTTCCCTTACCTTGGGTCTCATCATCAATGGAATCTT 961  
 QY 1066 CTGAGGACAGTTTCAAGACTTGTCTGCTCCCAATGTAGAGACACCACTATGTTCC 1125  
 DB 962 CTGAGGACAAATCCAAAGACTTTTCTGCTCCCAATGTGAGGACCACTATGTTCC 1021  
 QY 1126 AATTGTAG 1133  
 DB 1022 AGTTATAG 1029  
 RESULT 15  
 ADE31648  
 ID ADE31648 standard; DNA; 1029 BP.  
 XX  
 AC ADE31648;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human 6193 gene #SEQ ID 5.  
 XX  
 KW Antiarteriosclerotic; cardiatic; vasotropic; antiinflammatory;  
 KW thrombolytic; antiarrhythmic; antianaginal; hypotensive; gene therapy;  
 KW cardiovascular; disease; ischaemia; aortic bending;  
 KW vascular heart disease; endocarditis; atrial fibrillation; heart failure;  
 KW angina; cardiomyopathy; cardiac death; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003065984-A2.  
 XX  
 PD 14-AUG-2003.  
 XX  
 PF 29-JAN-2003; 2003WO-US002571.  
 XX  
 PR 01-FEB-2002; 2002US-0353224P.  
 PR 15-MAR-2002; 2002US-0364529P.  
 PR 19-APR-2002; 2002US-0373861P.  
 PR 29-APR-2002; 2002US-0376287P.  
 PR 12-JUN-2002; 2002US-038080P.  
 PR 24-JUN-2002; 2002US-039091P.  
 PR 03-JUL-2002; 2002US-0394130P.  
 PR 10-JUL-2002; 2002US-0394797P.  
 PR 21-AUG-2002; 2002US-0405450P.  
 PR 04-SEP-2002; 2002US-0408070P.  
 PR 06-NOV-2002; 2002US-0424300P.  
 PR 05-DEC-2002; 2002US-0431042P.  
 PR 05-DEC-2002; 2002US-0431079P.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Logan TJ, Chun M, Galvin KM, Healy A, Acton SL, Donaghue M;  
 PI Stagliano N, Perodin J, Rodrigue-Way A;  
 XX  
 DR WPI; 2003-731468/59.  
 DR P-PSDB; ADE31649.  
 XX  
 PT Identifying a compound capable of treating a cardiovascular disorder  
 PT (e.g. atherosclerosis) comprises assaying the ability of the compound to  
 PT modulate the expression or activity of e.g. 1682, 6169 or 6193  
 PT polypeptide or nucleic acid.  
 XX  
 PS Disclosure; SEQ ID NO 5; 328pp; English.



XX The invention relates to a method for identifying a compound capable of  
 CC treating a cardiovascular disorder. The present invention identifies the  
 CC differential expression of 1682, 6169, 6193, 14395, 29002, 33216,  
 CC 43726, 69292, 21656, 32427, 2402, 7747, 1720, 9151, 60491, 1371, 7077,  
 CC 33207, 1419, 18036, 16105, 38650, 14245, 58848, 1870, 25856, 32394, 3484,  
 CC 345, 9252, 9135, 10332, 18610, 8165, 2448, 2445, 84624, 84237, 8912,  
 CC 2868, 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2314, 51130,  
 CC 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474, 9792, 15400, 1452 or  
 CC 6585 genes in cardiovascular disease states. The methods are useful in  
 CC diagnosing, preventing and treating cardiovascular disorders, such as  
 CC atherosclerosis, cardiac hypertrophy, ischaemia, reperfusion injury,  
 CC stenosis, arterial inflammation, vascular wall remodeling, coronary  
 CC microembolism, tachycardia, bradycardia, pressure overload, aortic  
 CC bending, coronary artery ligation, vascular heart disease, valvular  
 CC disease, including but not limited to, valvular degeneration caused by  
 CC calcification, rheumatic heart disease, endocarditis, or complications of  
 CC artificial valves; atrial fibrillation, long-QT syndrome, congestive  
 CC heart failure, sinus node dysfunction, angina, heart failure,  
 CC hypertension, atrial fibrillation, atrial flutter, pericardial disease,  
 CC including but not limited to, pericardial effusion and pericarditis;  
 CC cardiomyopathies, e.g. dilated cardiomyopathy or idiopathic  
 CC cardiomyopathy, myocardial infarction, coronary artery disease, coronary  
 CC artery spasm, ischaemic disease, arrhythmia, sudden cardiac death and  
 CC cardiovascular developmental disorders. The methods may also be used for  
 CC identifying compounds that modulate cardiovascular disorders. Sequences  
 CC given in ADE31644-ADE31769 represent the genes and proteins that may be  
 CC regulated by a compound of the invention.

XX  
 SQ Sequence 1029 BP; 229 A; 275 C; 238 G; 287 T; 0 U; 0 Other;

Query Match 54.3%; Score 679.2; DB 9; Length 1029;  
 Best Local Similarity 81.2%; Pred. No. 6.4e-201;  
 Matches 786; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 166 AGGAGAACAAAGCTTCTCTAAAGTTCAGAGAGTCTTTTGGCCCTGTGTGACCTGGTAG 225  
 DB |||||  
 QY 226 TGTGTGCTTTGGAGTCTAGGAATCTCCCTGGTCTGATTATATACATTTCTACACAGA 285  
 DB |||||  
 DB 122 TGTGTGCTGTGTGTGGGAACTCTCTGGTCTGGTCAATCCATCTTCTACCATTA 181  
 QY 286 AGCTGAGGACTCTGACAGATGTTTCTCTGAATCTGCCCTGCTGACCTGGTGTGTTG 345  
 DB |||||  
 DB 182 AGTTGACAGGCTGACGAGTGTCTCTGGTGAACCTACCCCTGCTGACCTGGTGTGTTG 241  
 QY 346 TCTGTACTCTGCCCTTTGGGCTATGAGGACCTATGAGTGGTCTTTGGCACAGTCA 405  
 DB |||||  
 DB 242 TCTGCACCTGCCCCCTCTGGGCTATGACGAGCATCCATGAATGGTGTGGCCAGGTCA 301  
 QY 406 TGTCAAAACTCTCGAGGCATGTATCAATGAATCTTACGTGTCCATGTCTCACTCTCA 465  
 DB |||||  
 DB 302 TGTGACAGGCTACTGGGCATCTACATATTAACTTCTACACGTCCTATGCTCATCTCA 361  
 QY 466 CTGTGATCAGTGGATCGTTTCATTTAGTGGTCCAGGCTACCAAGGCTTCAACCGGC 525  
 DB |||||  
 DB 362 CTGTGATCACTGTGGATCGTTTCATTTAGTGGTGAAGGCCACCAAGGCTTCAACAGC 421  
 QY 526 AGGCTAAGTGGAGATCTGGGCCAAGTCAATTTGCTGCTATTTGGTGGTCTCCCTGT 585  
 DB |||||  
 DB 422 AAGCCAGAGGATGACCTGGGCAAGGTCCACAGCTTGTCTATCTGGGTGATATCCCTGC 481  
 QY 586 TGGTTTCTTTGCCACAGATCATCTATGGCCATGTTTCAAGATATTCAAGCTTATCTGTC 645  
 DB |||||  
 DB 482 TGGTTTCTTTGCCCAAAATATCTATGGCAATGTCTTTAATCTCGACAAGCTCATATGTG 541  
 QY 645 AGTACCAGAGTGGAGATATCCATATGTTCTGTTATACAGATGACTCTGGGGTCT 705  
 DB |||||  
 DB 542 GTTACCATGACGAGGAATTTCCATCTGGTCTTTCGCCACCCAGATGACACTGGGGTCT 601  
 QY 706 TCCTGCATTTGCTCACTATGATTCTGTCTACTCAGGCATTTATCAAGACCTTGTCTCATG 765  
 DB |||||

DB 602 TCTTGGCAGCTCAGCATGATTTGTCTGCTATTTCAGTCAATCAAAACACTGCTTCATG 661  
 QY 766 CTCGAAACTTCCAGAGACAAATCTCTAAAGATCATCTTCTGTAGTGGTGTGTTCC 825  
 DB |||||  
 DB 662 CTGAGGCTTCCAGAGACAGATCTCTAAAGATCATCTTCTGTGTGATGGCTGTGTTCC 721  
 QY 826 TGTGACCCAGACACACCTTCAACCTTGGCCATGTTAATCCAAAGTACAAGCTGGGAGTACT 885  
 DB |||||  
 DB 722 TGTGACCCAGATGCCCCCTTCAACCTCATGAAGTTCACTCGCACACACACTGGGAATACT 781  
 QY 886 ATACCATAACACAGCTTTAAGTATGCCATCGTAGTCAGAGGCTATAGCATACTTTCCGGG 945  
 DB |||||  
 DB 782 ATGCCATGACCACTTTCTACTACACCATCATGCTGACAGAGGCCATCGCATACCTGAGGG 841  
 QY 946 CTTCGCTTAACCTGTACTTTATGCGCTTTGTTGGCTTAAAGTTCGGAAGAACGTTCTGA 1005  
 DB |||||  
 DB 842 CTTGCTTAAACCTGTGCTCTATGCGCTTTGTACGCTGAAAGTTTCGAAAGAACTTCTGA 901  
 QY 1006 AACTTATGAAGGATATCGGCTGCTCTCTCACCTGGGAGTCTCAAGTCAATGGAAGTCTT 1065  
 DB |||||  
 DB 902 AACTTGTGAGGACATTTGGTTGCTCCCTTACCTTGGGCTCTCACATCAATGGAATCTT 961  
 QY 1066 CTGAGGACATTTCCAGACTTGTGCTTCCCTCCACAAATGTAGAGACACACAGTATGTTCC 1125  
 DB |||||  
 DB 962 CTGAGGACATTTCCAGACTTGTGCTTCCCTCCCAATGTGGAGGCCACACAGCATGTTCC 1021  
 QY 1126 AATTGTAG 1133  
 DB |||||  
 DB 1022 AGTTATAG 1029

Search completed: September 29, 2004, 00:45:43  
 Job time : 527 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2004, 00:32:09 ; Search time 3116 Seconds

(without alignments)  
11988.956 Million cell updates/sec

Title: US-10-603-141-1

Perfect score: 1251

Sequence: 1 Gaattcgacagaggcaga.....gatatcgaggagcctgac 1251

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estm:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hnc1:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_nam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gsl1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1231.4	98.4	1791	11 AK052901	AK052901 Mus muscu
2	1029	82.3	1029	29 AY399296	AY399296 Mus muscu
3	859	68.7	991	13 BY746127	BY746127 BY746127
4	679.2	54.3	1029	29 AY399294	AY399294 Homo sapi

5	679.2	54.3	1029	29	AY399295	AY399295 Pan trogl
6	662	52.9	672	13	BY746487	BY746487 BY746487
7	658.2	52.6	663	14	CF912719	CF912719 A0636C03-
8	647	51.7	647	14	CF911440	CF911440 A0615D03-
9	631	50.4	631	14	CF913009	CF913009 A0641F09-
10	627	50.1	627	14	CF911103	CF911103 A0610D08-
11	626.4	50.1	628	14	CF912382	CF912382 A0631A05-
12	617.4	49.4	619	14	CA574451	CA574451 K0616E05-
13	616	49.2	626	10	BB662193	BB662193 BB662193
14	610	48.8	610	14	CF910757	CF910757 A0604D11-
15	599.4	47.9	601	14	CF913351	CF913351 A0647A08-
16	596.2	47.7	601	14	CF911739	CF911739 A0620D02-
17	594.4	47.5	596	14	CA575926	CA575926 K0637E06-
18	587.5	47.0	594	13	BQ562713	BQ562713 H4077E09-
19	584.4	46.7	586	14	CF911576	CF911576 A0617F05-
20	575.6	46.0	773	12	BG363872	BG363872 G0282855
21	572.6	45.8	579	14	CA574155	CA574155 K0611B04-
22	567.4	45.4	569	14	CA576578	CA576578 K0646H04-
23	558.4	44.6	560	14	CF912084	CF912084 A0626A08-
24	555.8	44.4	629	14	CF910917	CF910917 A0606H11-
25	549.2	43.9	554	14	CF912157	CF912157 A0627C05-
26	549	43.9	608	10	BB660928	BB660928 BB660928
27	539	43.1	659	10	BB630116	BB630116 BB630116
28	514	41.1	884	12	BG963267	BG963267 G0282836
29	504.4	40.3	506	14	CA575464	CA575464 K0631A11-
30	492	39.3	514	14	CF911881	CF911881 A0622G04-
31	484	38.7	484	14	CA575959	CA575959 K0637H11-
32	482.4	38.6	484	14	CA573379	CA573379 K0600C07-
33	482.4	38.6	484	14	CA574776	CA574776 K0621A04-
34	462	36.9	462	14	CF911521	CF911521 A0616F03-
35	452.8	36.2	467	14	CA574254	CA574254 K0613E05-
36	438.4	35.0	440	14	CA576368	CA576268 K0643C07-
37	394	31.5	416	14	CA576715	CA576715 K0649B10-
38	391.6	31.3	1060	13	BX398345	BX398345 BX398345
39	385.8	30.8	1201	13	BX398107	BX398107 BX398107
40	381.2	30.5	650	12	BI837704	BI837704 G03086564
41	366.6	29.3	622	9	AI648943	AI648943 uk333b08.x
42	357.4	28.6	368	13	BY214569	BY214569 BY214569
43	352	28.1	679	28	BZ263559	BZ263559 CH230-254
44	347	27.7	356	13	BY221709	BY221709 BY221709
45	345	27.6	354	13	BY221071	BY221071 BY221071

## ALIGNMENTS

RESULT 1  
AK052901  
LOCUS  
DEFINITION  
Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched library, clone:D830030f09 product:chemokine (C-X-C) receptor 6, full insert sequence.  
1791 bp mRNA linear HTC 20-SEP-2003  
ACCESSION  
AK052901  
VERSION  
AK052901.1 GI:26343078  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE  
1  
Carninci, P. and Hayashizaki, Y.  
AUTHORS  
High-efficiency full-length cDNA cloning  
TITLE  
JOURNAL  
Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE  
99279253  
PUBMED  
10349636  
REFERENCE  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL  
Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE  
20499374  
PUBMED  
11042159

REFERENCE  
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishire, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

## TITLE

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL  
MEDLINE  
PUBMED

Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

REFERENCE  
AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 403, 685-690 (2001)

## TITLE

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection

REFERENCE  
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

## TITLE

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL  
MEDLINE  
PUBMED

Nature 420, 563-573 (2002)  
6 (bases 1 to 1791)

REFERENCE  
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, K., Saichon, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

## TITLE

Direct Submission  
Submitted (15-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES  
source

Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers

1. 1791  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:D830030F09"  
/db\_xref="MGI:2423199"  
/db\_xref="taxon:10090"  
/clone="D830030F09"  
/tissue\_type="heart"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="16 days neonate"  
73. 1128  
/note="unnamed protein product; chemokine (C-X-C) receptor  
6 (MGI:MGI:1934582, GB|NM\_030712, evidence: BLASTN, 99%, match=1788)  
putative"  
/codon\_start=1  
/protein\_id="BAC35196.1"  
/db\_xref="GI:26343079"

## CDS

1128  
/note="unnamed protein product; chemokine (C-X-C) receptor  
6 (MGI:MGI:1934582, GB|NM\_030712, evidence: BLASTN, 99%, match=1788)  
putative"  
/codon\_start=1  
/protein\_id="BAC35196.1"  
/db\_xref="GI:26343079"

/translation="MDDGHQBSALYDGHYEGDFLNNSSDNQENKRFLEKFEVFLP  
CVLVVVFVGLLGNVLIIIVFYKRLTLTDVLLNPLADLVFVCTLPFWVAGTV  
EWFTVWCKTLRGMYTNFVYVSMLTLCITVDREIVVQVATKAFNQAQKWIWGOVI  
CLLIMVSLVSLPOLIYGHVODIDKLIQVHSEISTMVLAVIOMTIGFELPLTML  
CVSGLIKLLHARNFQKHKSUKIIFLVVAVFLITQTFNLANLIQSSWEYTTTSRK  
YAVVTEIAIYFRACINLPIYAFVGLKFRKRWKMLMDIGICLSHLGVSSQWKSSEDS  
KTCASASHNVETTSFQL"

## ORIGIN

Query Match 98.4%; Score 1231.4; DB 11; Length 1791;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1232; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	15	GGCAGACCTTTAGTGAGCACACTTCACTCTGGAACAAAGCTACTGGGCTTCTCTTCTGAT	74
DB	10	GGCAGACCTTTAGTGAGCACACTTCACTCTGGAACAAAGCTACTGGGCTTCTCTTCTGAT	69
QY	75	GCATGGATGATGGGCATCAAGATCAGCTCTGTACGATGGGCATCGAGGGAGATTTC	134
DB	70	GCATGGATGATGGGCATCAAGATCAGCTCTGTACGATGGGCATCGAGGGAGATTTC	129
QY	135	TGGCTCTTCAACAATCCAGTGATAACAGCCAGGAGAACAAACGCTTCTCTAAAGTCAAG	194
DB	130	TGGCTCTTCAACAATCCAGTGATAACAGCCAGGAGAACAAACGCTTCTCTAAAGTCAAG	189
QY	195	GAGGCTTTTGGCCCTGTGTACTGTGTAGTGTGTCTTTGGACTGCTAGGAACTCC	254
DB	190	GAGGCTTTTGGCCCTGTGTACTGTGTAGTGTGTCTTTGGACTGCTAGGAACTCC	249
QY	255	CTGGTCTTGATTATATACATTTTCTACAGAGCTGAGGACTCTGACAGATGTCTTCTG	314
DB	250	CTGGTCTTGATTATATACATTTTCTACAGAGCTGAGGACTCTGACAGATGTCTTCTG	309
QY	315	CTGAATTTGGCCCTGGCTGACCTGGTGTGTCTGTACTCTGCCCTTTTGGCCCTATGCA	374
DB	310	CTGAATTTGGCCCTGGCTGACCTGGTGTGTCTGTACTCTGCCCTTTTGGCCCTATGCA	369
QY	375	GGCAGCTATGATGGTCTTTGGCAGCATGATGCAAACTCTTCGAGCATCTATACA	434
DB	370	GGCAGCTATGATGGTCTTTGGCAGCATGATGCAAACTCTTCGAGCATCTATACA	429
QY	435	ATGAATCTTACGTGTCCATCTCACTCTACCTGCATCAGTCGATCGTTTCATTGTA	494
DB	430	ATGAATCTTACGTGTCCATCTCACTCTACCTGCATCAGTCGATCGTTTCATTGTA	489
QY	495	GTGGTCCAGGCTACAGGCCCTTCAACGGCAGGCTAAGTGAAGATCTGGGCCAAGTC	554
DB	490	GTGGTCCAGGCTACAGGCCCTTCAACGGCAGGCTAAGTGAAGATCTGGGCCAAGTC	549
QY	555	ATTGCTTGTCTCATTTGGGTGCTCTCCCTGTGTGTTTCTTTGCCACAGATCATCTATGC	614
DB	550	ATTGCTTGTCTCATTTGGGTGCTCTCCCTGTGTGTTTCTTTGCCACAGATCATCTATGC	609
QY	615	CATGTTCAAGATATGACAGCTTATCTGTACGTACACAGTGGAGAGATATCCACTATG	674
DB	610	CATGTTCAAGATATGACAGCTTATCTGTACGTACACAGTGGAGAGATATCCACTATG	669
QY	675	GTTCTTGTATACAGATGACTCTGGGTCTTCTCTGCTGCTTCTTCTGATTTCTGTC	734
DB	670	GTTCTTGTATACAGATGACTCTGGGTCTTCTCTGCTGCTTCTTCTGATTTCTGTC	729
QY	735	TACTCAGGATATCAAGACCTTCTGCTCATGCTCGAACTTCCAGAAAGCAAAATCTCTA	794
DB	730	TACTCAGGATATCAAGACCTTCTGCTCATGCTCGAACTTCCAGAAAGCAAAATCTCTA	789
QY	795	AAGATCATCTTCTTGTAGTGGTGTCTCTGCTGACCCAGACACCCCTTCAACCTTGGC	854
DB	790	AAGATCATCTTCTTGTAGTGGTGTCTCTGCTGACCCAGACACCCCTTCAACCTTGGC	849
QY	855	ATGTTAATCCAAAGTCAAGCTGGAGTACTATACCAATAACCAAGCTTTAAGTAGTGCATC	914
DB	850	ATGTTAATCCAAAGTCAAGCTGGAGTACTATACCAATAACCAAGCTTTAAGTAGTGCATC	909

```

QY 915 GTAGTGACAGAGGCTATAGCATACCTTCGGGCTTGCCTTAACCCCTGACTTTATGCGCTTT 974
Db 910 GTAGTGACAGAGGCTATAGCATACCTTCGGGCTTGCCTTAACCCCTGACTTTATGCGCTTT 969
QY 975 GTTGGCTTAAAGTTCGCGAAGAACTCTGGAACCTTATGAAGGATATCGGCTGCTCTCT 1034
Db 970 GTTGGCTTAAAGTTCGCGAAGAACTCTGGAACCTTATGAAGGATATCGGCTGCTCTCT 1029
QY 1035 CACCTGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTCCTCAAGACTTGTTCGCC 1094
Db 1030 CACCTGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTCCTCAAGACTTGTTCGCC 1089
QY 1095 TCCCAATATGAGACACACAGATGTTCCAAATGTAGTAGGCTTGGCCACACTTAGAG 1154
Db 1090 TCCCAATATGAGACACACAGATGTTCCAAATGTAGTAGGCTTGGCCACACTTAGAG 1149
QY 1155 AAGTTAATAACAGAACTTCTAGGACATGCTGTATCATTTGGATGCAACAAGAAAGCTT 1214
Db 1150 AAGTTAATAACAGAACTTCTAGGACATGCTGTATCATTTGGATGCAACAAGAAAGCTT 1209
QY 1215 TCGTTATAGCATGTGGAGTATCATGGAAGCC 1247
Db 1210 TCGTTATAGCATGTGGAGTATCATGGAAGTC 1242

RESULT 2
AY399296 1029 bp DNA linear GSS 15-DEC-2003
LOCUS Mus musculus CXCR6 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY399296
VERSION AY399296.1 GI:39755285
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1029)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1029)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submision
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1029
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1029
/gene="CXCR6"
/locus_tag="HMC0176"
ORIGIN
Query Match 82.3%; Score 1029; DB 29; Length 1029;
Best Local Similarity 100.0%; Freq. No. 7.1e-293;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 CTGTAGATGGGCACTACGAGGAGATTTCTGGCTCTTCAACAATTCAGTGATACAGC 164
Db 1 CTGTAGATGGGCACTACGAGGAGATTTCTGGCTCTTCAACAATTCAGTGATACAGC 60

```

```

QY 165 CAGAGAAACAAGCTTCTTAAAGTTCAAGAGAGTCTTTTTTGCCTGTGTGACCTGGTA 224
Db 61 CAGAGAAACAAGCTTCTTAAAGTTCAAGAGAGTCTTTTTTGCCTGTGTGACCTGGTA 120
QY 225 GTCTTTGTCTTTGGACTGTAGGAAACTCCCTGGTCTCTGATTATATACATTTTCTACACAG 284
Db 121 GTCTTTGTCTTTGGACTGTAGGAAACTCCCTGGTCTCTGATTATATACATTTTCTACACAG 180
QY 285 AAGCTGAGGACTCTGACAGATGTTTTCTGCTGAACCTTGCCTTGGCTGCTGACCTGGTGT 344
Db 181 AAGCTGAGGACTCTGACAGATGTTTTCTGCTGAACCTTGCCTTGGCTGCTGACCTGGTGT 240
QY 345 GTCTGTACTCTGCCCTTTGGGCTATGAGGACACCTATGAGTGGTCTTTTGGSCACAGTC 404
Db 241 GTCTGTACTCTGCCCTTTGGGCTATGAGGACACCTATGAGTGGTCTTTTGGSCACAGTC 300
QY 405 ATGTGCAAAACTCTTCGAGGCATGTATACAATGAACCTTCTAGGTGCCATGCTCATCTTC 464
Db 301 ATGTGCAAAACTCTTCGAGGCATGTATACAATGAACCTTCTAGGTGCCATGCTCATCTTC 360
QY 465 ACCTGCATCACAGTGGATCGTTTTCATTTAGTGTCCAGGCTACCAAGGCTTCAACCCG 524
Db 361 ACCTGCATCACAGTGGATCGTTTTCATTTAGTGTCCAGGCTACCAAGGCTTCAACCCG 420
QY 525 CAGGCTAAGTGAAGATCTGGGGCCAAAGTCATTTGCTTGTCTCATTTGGGTGGTCTCCCTG 584
Db 421 CAGGCTAAGTGAAGATCTGGGGCCAAAGTCATTTGCTTGTCTCATTTGGGTGGTCTCCCTG 480
QY 585 TTGGTTCTTTGGCCACAGATCATCTATGGCCATGTTCAAGATATTGACAAGCTTATCTGT 644
Db 481 TTGGTTCTTTGGCCACAGATCATCTATGGCCATGTTCAAGATATTGACAAGCTTATCTGT 540
QY 645 CAGTACCACAGTCAGGAGATCCACATGTTGTTCTTGTATTACAGATGACTCTGGGGTTC 704
Db 541 CAGTACCACAGTCAGGAGATCCACATGTTGTTCTTGTATTACAGATGACTCTGGGGTTC 600
QY 705 TTCTGCAATTTGCTACTATGATTTCTGTGCTACTCAGGCAATTAAGACCTTCTTCAT 764
Db 601 TTCTGCAATTTGCTACTATGATTTCTGTGCTACTCAGGCAATTAAGACCTTCTTCAT 660
QY 765 GCTCGAAGCTTCAGAGACAAATCTTAAGATCATCTTCCCTGTAGTGGCTGTGTTTC 824
Db 661 GCTCGAAGCTTCAGAGACAAATCTTAAGATCATCTTCCCTGTAGTGGCTGTGTTTC 720
QY 825 CTGCTGACCCAGACACCCCTTCAACCTTGCATGTTAATCCAAAGTACAAGCTGGAGTAC 884
Db 721 CTGCTGACCCAGACACCCCTTCAACCTTGCATGTTAATCCAAAGTACAAGCTGGAGTAC 780
QY 885 TATACCAATACAGCTTTAAGTATGCCATGCTAGTGACAGAGCTATAGCATACTTTCGG 944
Db 781 TATACCAATACAGCTTTAAGTATGCCATGCTAGTGACAGAGCTATAGCATACTTTCGG 840
QY 945 GCTTGCCTTAAACCTGTACTTTATGCTTTGTTGGCTTAAAGTTCGGAAGAACGTCGG 1004
Db 841 GCTTGCCTTAAACCTGTACTTTATGCTTTGTTGGCTTAAAGTTCGGAAGAACGTCGG 900
QY 1005 AAACCTTAAAGGATATCGGCTGCTCTCTCACTGGAGTCTCAAGTCAATGGAAGTCT 1064
Db 901 AAACCTTAAAGGATATCGGCTGCTCTCTCACTGGAGTCTCAAGTCAATGGAAGTCT 960
QY 1065 TCTGAGCACAGTTCCTAGACTTGTTCCTGCCCTCCACAAATGTAGAGACCAAGTATGTTTC 1124
Db 961 TCTGAGCACAGTTCCTAGACTTGTTCCTGCCCTCCACAAATGTAGAGACCAAGTATGTTTC 1020
QY 1125 CAATTGTAG 1133
Db 1021 CAATTGTAG 1029

```

RESULT 3  
BY746127  
LOCUS

991 bp mRNA linear EST 17-DEC-2002

DEFINITION BY746127 RIKEN full-length enriched, 2 days neonate thymus thymic cells (NOD) Mus musculus cDNA clone E43004020 5', mRNA sequence.

ACCESSION BY746127

KEYWORDS BY746127.1 GI:27173642

SOURCE EST. Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 931)

AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamana, K., Kiyosawa, H., Iagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Bruscia, V., Chotha, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M., King, B. D., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C. A., Setou, M., Shimada, K., Sultana, R., Tacken, L., Taylor, M. S., Teasdale, R. D., Tomita, M., Velazquez, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wyshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 22354683

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsr.riken.go.jp  
URL: http://genome.gsc.riken.go.jp/  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Hirozane, T., Hori, F., Fukuda, S., Hashizume, W., Hayashida, K., Hayashida, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/WRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES  
Location/Qualifiers  
1..991  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="NOD"  
/db\_xref="taxon:10090"  
/clone="E43004020"  
/tissue\_type="thymus"  
/cell\_type="thymic cells"  
/clone\_lib="RIKEN full-length enriched, 2 days neonate thymus thymic cells (NOD)"

ORIGIN  
Query Match 68.7%; Score 859; DB 13; Length 991;  
Best Local Similarity 95.1%; Pred. No. 1.6e-242;  
Matches 938; Conservative 0; Mismatches 43; Indels 5; Gaps 5;

QY 11 CGAGGGCAGACCTTTAGTGAGCACACTTCACTCTGGACAAAGCTACTGGGCTTCTCTTC 70  
DB 7 CGGAGGCAGACCTTTAGTGAGCACACTTCACTCTGGACAAAGCTACTGGGCTTCTCTTC 66  
QY 71 TGATGCCATGATGATGGGCATCAAGAGTCAGCTCTGTACGATGGCAGCTACGAGGAGA 130  
DB 67 TGATGCCATGATGATGGGCATCAAGAGTCAGCTCTGTACGATGGCAGCTACGAGGAGA 126  
QY 131 TTTCTGGCTCTTCAACAAATTCAGTGATTAACAGCCAGAGAGAACAAAGCTCTCTAAAGTT 190  
DB 127 TTTCTGGCTCTTCAACAAATTCAGTGATTAACAGCCAGAGAGAACAAAGCTCTCTAAAGTT 186  
QY 191 CAAGAGAGCTTTTGGCCCTGTGTACTCTGTGTAGTGTGTCTTGTGAGCTGTAGGAAA 250  
DB 187 CAAGAGAGCTTTTGGCCCTGTGTACTCTGTGTAGTGTGTCTTGTGAGCTGTAGGAAA 246  
QY 251 CTCCTGGTTCTGATTATATATATTTTACCAAGCTGAGGACTCTGACAGATGTTT 310  
DB 247 CTCCTGGTTCTGATTATATATATTTTACCAAGCTGAGGACTCTGACAGATGTTT 306  
QY 311 TCTGTGAACCTGCCCTGGCTGACCTGGTGTGTGTCTGTACTCTGCGCTTTGGGCTA 370  
DB 307 TCTGTGAACCTGCCCTGGCTGACCTGGTGTGTGTCTGTACTCTGCGCTTTGGGCTA 366  
QY 371 TGCAGGCACCTATGAGTGGGTCTTTTGGCACAGTCAATGTGCAAAACTCTTCGAGGCATGTA 430  
DB 367 TGCAGGCACCTATGAGTGGGTCTTTTGGCACAGTCAATGTGCAAAACTCTTCGAGGCATGTA 426  
QY 431 TACATGAACTTCTACGTGTCTGCTCACTCTCACTGATCACTGATGATGCTTTTTCAT 490  
DB 427 TACATGAACTTCTACGTGTCTGCTCACTCTCACTGATCACTGATGATGCTTTTTCAT 486  
QY 491 TGTAGTGTCTCCAGGCTACCAAGGCTTCAACCGGCGAGGCTAAAGTGAAGATCTTGGGGCCA 550  
DB 487 TGTAGTGTCTCCAGGCTACCAAGGCTTCAACCGGCGAGGCTAAAGTGAAGATCTTGGGGCCA 546  
QY 551 AGTCATTTGCTTGTCTCAATTTGGGTGTCTCTCTGTTTGGTCTTTTGGCCACAGATCATCTA 610  
DB 547 AGTCATTTGCTTGTCTCAATTTGGGTGTCTCTCTGTTTGGTCTTTTGGCCACAGATCATCTA 606  
QY 611 TGGCCATTTTCAAGATATTGCAAGCTTATCTGTCAGTACCACAGTGAAGATATCCAC 670  
DB 607 TGGCCATTTTCAAGATATTGCAAGCTTATCTGTCAGTACCACAGTGAAGATATCCAC 666  
QY 671 TATGGTTCTTTTATACAGATGACTCTGGGTTCTTCTGCTTCTTCTGCTTCTCTATGATCT 730  
DB 667 TATGGTTCTTTTATACAGATGACTCTGGGTTCTTCTGCTTCTTCTGCTTCTCTATGATCT 726







MEDLINE  
PUSHED  
COMMENT

22354683  
12466851  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/  
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Hirozane,T., Hori,F.,  
Ikumori,S., Hashizume,W., Hayashida,K., Hirozane,T., Kawai,J., Koijima,Y.,  
Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,  
Nomura,K., Numasaki,D., Ohno,M., Ohsato,N., Saito,R., Sakagami,N.,  
Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,  
Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.  
Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
Prepare mouse tissues.  
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics  
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome  
Trust/MRC building Addenbrookes Hospital Cambridge) whose  
assistance we gratefully acknowledge.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.

FEATURES  
source  
1..672  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="NOD"  
/db\_xref="taxon:10090"  
/clone="E430014H02"  
/tissue\_type="thymus"  
/cell\_type="thymic cells"  
/clone\_lib="RIKEN full-length enriched, 2 days neonate  
thymus thymic cells (NOD)"

ORIGIN

Query Match 52.9%; Score 662; DB 13; Length 672;  
Best Local Similarity 100.0%; Pred.No. 2.7e-184;  
Matches 682; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 15 GCAGACCTTTAGTGAGCACACTTCACCTCTCGGAACAAGTAGCTGTGGCTTCTCTTCGTAT 74  
Db 11 GCAGACCTTTAGTGAGCACACTTCACCTCTCGGAACAAGTAGCTGTGGCTTCTCTTCGTAT 70  
QY 75 GCCATGGATGATGGGCATCAGAGTCAGCTCTGTACGATGGGCATCCAGGGAGATTTC 134  
Db 71 GCCATGGATGATGGGCATCAGAGTCAGCTCTGTACGATGGGCATCCAGGGAGATTTC 130  
QY 135 TGCGCTCTTCAACAATTCAGTGATTAACAGCCAGGAGAACAACGCTTCCTAAAGTTTCAAG 194  
Db 131 TGCGCTCTTCAACAATTCAGTGATTAACAGCCAGGAGAACAACGCTTCCTAAAGTTTCAAG 190  
QY 195 GAGGTCTTTTGGCCCTGTGTGTACTGTAGTGTGTTTGTCTTTGGAGCTGTAGAAAACCTCC 254  
Db 191 GAGGTCTTTTGGCCCTGTGTGTACTGTAGTGTGTTTGTCTTTGGAGCTGTAGAAAACCTCC 250



average insert size is about 2.2 kb. The library was constructed by Yulan Piao."

## ORIGIN

Query Match 51.7%; Score 647; DB 14; Length 647;  
Best Local Similarity 100.0%; Pred. No. 7.3e-180;  
Matches 647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 ATGGGCACTACGAGGAGATTTCTGGCTCTCAACAATTCAGATGATACAGCCAGGAGA 171  
Db 1 ATGGGCACTACGAGGAGATTTCTGGCTCTCAACAATTCAGATGATACAGCCAGGAGA 60

QY 172 ACAACGCTTCTCAAGTTCAGGAGGCTTTTTCGCCCTGTGTGTACCTGGTGTAGTGTG 231  
Db 61 ACAACGCTTCTCAAGTTCAGGAGGCTTTTTCGCCCTGTGTGTACCTGGTGTAGTGTG 120

QY 232 TCTTTGGACTGTAGAACTCCCTGGTCTGATATATACATTTCTACCAAGAGCTGA 291  
Db 121 TCTTTGGACTGTAGAACTCCCTGGTCTGATATATACATTTCTACCAAGAGCTGA 180

QY 292 GGACTCTGACAGATGTTTCTGCTGAACTTGCCTCGCTGACCTGGTGTGTGTCTGTA 351  
Db 181 GGACTCTGACAGATGTTTCTGCTGAACTTGCCTCGCTGACCTGGTGTGTGTCTGTA 240

QY 352 CTCTGCCCTTTTGGCCCTATGAGGACACTATGAGTGGGTCTTTGGCCAGATCATGTGCA 411  
Db 241 CTCTGCCCTTTTGGCCCTATGAGGACACTATGAGTGGGTCTTTGGCCAGATCATGTGCA 300

QY 412 AAACCTCTCGAGCATGTATACAATGAACTTCACTGTCCATGCTCACTCTCACTGCA 471  
Db 301 AAACCTCTCGAGCATGTATACAATGAACTTCACTGTCCATGCTCACTCTCACTGCA 360

QY 472 TCACAGTGGATCGTTTCATTTAGTGGTCCAGGCTACCAAGCCCTTCAACCGGAGGCTA 531  
Db 361 TCACAGTGGATCGTTTCATTTAGTGGTCCAGGCTACCAAGCCCTTCAACCGGAGGCTA 420

QY 532 AGTGAAGATCTGGGGCAAGTCATTTGCTCTCATTTGGTGGTCTCCCTGTGTGTTT 591  
Db 421 AGTGAAGATCTGGGGCAAGTCATTTGCTCTCATTTGGTGGTCTCCCTGTGTGTTT 480

QY 592 CTTTCCCAACAGATCATCTATGGCCATGTTCAGATATTCACAAGCTATCTGTCACTGACC 651  
Db 481 CTTTCCCAACAGATCATCTATGGCCATGTTCAGATATTCACAAGCTATCTGTCACTGACC 540

QY 652 ACAGTGAGGAGATATCCATATGTTCTGTTATACAGATGACTCTGGGGTCTTCTCTGC 711  
Db 541 ACAGTGAGGAGATATCCATATGTTCTGTTATACAGATGACTCTGGGGTCTTCTCTGC 600

QY 712 CATTTCTCACTATGATTTCTGTCTACTCAGGCATTTATCAAGACCTTG 758  
Db 601 CATTTCTCACTATGATTTCTGTCTACTCAGGCATTTATCAAGACCTTG 647

RESULT 9  
CF913009  
LOCUS  
DEFINITION  
A0641F09-5 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)  
CDNA Library (Long 1) Mus musculus cDNA clone NIA:A0641F09  
IMAGE:30751556 5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)

REFERENCE  
AUTHORS  
Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.  
TITLE  
Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method

JOURNAL  
GENOME RES. 11 (9), 1553-1558 (2001)  
21429098

PUBMED  
COMMENT

11544199  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@lgsun.grc.nia.nih.gov  
Plate: A0641 row: F column: 09  
Seq primer: M13 Reverse  
High quality sequence stop: 631  
POLYA=No.

FEATURES  
source

Location/Qualifiers  
1. .631  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6NCr"  
/db\_xref="niaEST:A0641F09-5"  
/db\_xref="taxon:10090"  
/clone="NIA:A0641F09 IMAGE:30751556"  
/dev\_stage="Age -10 weeks old"  
/lab\_host="DH10B"  
/clone\_lib="NIA Mouse Hematopoietic Stem Cell  
(Lin-/c-Kit-/Sca-1+) cDNA Library (Long 1)"  
/notes="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;  
Site 2: NotI; Mouse cDNA project by the Laboratory of  
Genetics, National Institute on Aging (NIA), Intramural  
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).  
This is a long-transcript enriched cDNA library [Ref.  
Genome Res. 11: 1553-1558 (2001). PMID: 11544199]. Total  
RNAs were obtained from Drs. Dennis Taub, Dan Longo  
(National Institute on Aging, USA), Jonathan Keller  
(National Cancer Institute, USA). Double-stranded cDNAs  
were synthesized with an oligo(dT) primer [Invitrogen:  
5'-pGACTAGTCTAGATCGAGCGCGCCCTTTTCTTTT-3'] from  
1.1 ug of total RNA, treated with T4 DNA polymerase, and  
purified by ethanol-precipitation. The cDNAs were ligated  
to Loner-linker LL-Sal4, purified by phenol/chloroform, and  
separated from free linkers by Centricon 100. Then, the  
cDNAs were amplified by long-range high fidelity PCR using  
Ex Taq polymerase (Takara) with a primer Sal4-S. The  
products were purified by phenol/chloroform and Centricon  
100. The cDNAs were digested with SalI and NotI enzymes  
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid  
vector. The DH10B E. coli host was transformed with the  
ligation mixture by the standard chemical method. The  
average insert size is about 2.2 kb. The library was  
constructed by Yulan Piao."

## ORIGIN

Query Match 50.4%; Score 631; DB 14; Length 631;  
Best Local Similarity 100.0%; Pred. No. 3.9e-175;  
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 ATGGGCACTACGAGGAGATTTCTGGCTCTTCAACAATTCAGTGTATAACAGCCAGGAGA 171  
Db 1 ATGGGCACTACGAGGAGATTTCTGGCTCTTCAACAATTCAGTGTATAACAGCCAGGAGA 60

QY 172 ACAACGCTTCTCAAGTTCAGGAGGCTTTTTCGCCCTGTGTGTACCTGGTGTAGTGTG 231  
Db 61 ACAACGCTTCTCAAGTTCAGGAGGCTTTTTCGCCCTGTGTGTACCTGGTGTAGTGTG 120

QY 232 TCTTTGGACTGTAGAACTCCCTGGTCTGATATATACATTTTCTACCAAGAGCTGA 291  
Db 121 TCTTTGGACTGTAGAACTCCCTGGTCTGATATATACATTTTCTACCAAGAGCTGA 180

QY 292 GGACTCTGACAGATGTTTCTGCTGAATTCGCCCTCGCTGACCTGGTGTGTGTCTGTA 351  
Db 181 GGACTCTGACAGATGTTTCTGCTGAATTCGCCCTCGCTGACCTGGTGTGTGTCTGTA 240

QY 352 CTCTGCCCTTTTGGCCCTATGAGGACACTATGAGTGGGTCTTTGGCCAGATCATGTGCA 411  
Db 241 CTCTGCCCTTTTGGCCCTATGAGGACACTATGAGTGGGTCTTTGGCCAGATCATGTGCA 300

QY 412 AAACCTCTCGAGCATGTATACAATGAACTTCTACCTGTCCATGCTCACTCTCACTGCA 471

301 AATCTTCAGGAGATGATACATGAATGACCTTACGTTGCTCATGCTCACTCTCACCTGCA 360  
 472 TCACAGTGGATCGTTTCATTTGATGTGTCAGGTCACCAAGGCTTCAACCGCAGGCTA 531  
 361 TCACAGTGGATCGTTTCATTTGATGTGTCAGGTCACCAAGGCTTCAACCGCAGGCTA 420  
 532 AGTGGAGAGATCTGGGGCCAAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTCTTGTTT 591  
 421 AGTGGAGAGATCTGGGGCCAAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTCTTGTTT 480  
 592 CTTTGGCCACAGATCATCTATGGCCATGTTCAAGATATTGAAGCTTATCTGTCATGACC 651  
 481 CTTTGGCCACAGATCATCTATGGCCATGTTCAAGATATTGAAGCTTATCTGTCATGACC 540  
 652 ACAGTGGAGAGATCTGGGGCCAAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTCTTGTTT 711  
 541 ACAGTGGAGAGATCTGGGGCCAAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTCTTGTTT 600  
 712 CATTGCTCACTATGATCTGCTGCTACTCAGG 742  
 601 CATTGCTCACTATGATCTGCTGCTACTCAGG 631

RESULT 10  
 CF911103  
 LOCUS  
 DEFINITION A0610D08-5 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)  
 IMAGE:30748555 5', mRNA sequence.  
 CF911103  
 VERSION CF911103.1 GI:38182307  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1. (bases 1 to 627)  
 Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.  
 Construction of long-transcript enriched cDNA libraries from  
 submicrogram amounts of total RNAs by a universal PCR amplification  
 method  
 Genome Res. 11 (9), 1553-1558 (2001)  
 21429098  
 11544199  
 Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@lgsun.grc.nia.nih.gov  
 Plate: A0610 row: D column: 08  
 Seq primer: M13 Reverse  
 High quality sequence stop: 627  
 POLYA=No.

FEATURES  
 source  
 1. 627  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6NCr"  
 /db\_xref="nlaEST:A0610D08-5"  
 /db\_xref="taxon:10090"  
 /clone="NIA:A0610D08 IMAGE:30748555"  
 /dev\_stage="Age -10 weeks old"  
 /lab\_host="DH10B"  
 /clone\_lib="NIA Mouse Hematopoietic Stem Cell  
 (Lin-/c-Kit-/Sca-1+) cDNA Library (Long 1)"  
 /note="vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;  
 Site 2: NotI; Mouse cDNA project by the Laboratory of  
 Genetics, National Institute on Aging (NIA), Intramural  
 Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).  
 This is a long-transcript enriched cDNA library (Ref.  
 Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total  
 RNAs were obtained from Drs. Dennis Taub, Dan Longo

(National Institute on Aging, USA), Jonathan Keller  
 (National Cancer Institute, USA). Double-stranded cDNAs  
 were synthesized with an Oligo(dT) primer (Invitrogen:  
 5'-TGACTGTTTCAGTACGAGCGCGCCCTTTTCTTTT-3') from  
 1.1 ug of total RNA, treated with T4 DNA polymerase, and  
 purified by ethanol-precipitation. The cDNAs were ligated  
 to lone-linker LL-Sal4, purified by phenol/chloroform, and  
 separated from free linkers by Centricon 100. Then, the  
 cDNAs were amplified by long-range high fidelity PCR using  
 Ex Taq polymerase (Takara) with a primer Sal4-S. The  
 products were purified by phenol/chloroform and Centricon  
 100. The cDNAs were digested with SalI and NotI enzymes  
 and cloned into SalI/NotI site of pCMV-SPORT6 plasmid  
 vector. The DH10B E. coli host was transformed with the  
 ligation mixture by the standard chemical method. The  
 average insert size is about 2.2 kb. The library was  
 constructed by Yulan Piao."

## ORIGIN

Query Match 50.1%; Score 627; DB 14; Length 627;  
 Best Local Similarity 100.0%; Pred. No. 6e-174;  
 Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 112 ATGGGCACCTACGAGGAGATTTCTGGCTCTTCAACAATTCAGTGATAACAGCCAGGAGA 171  
 Db 1 ATGGGCACCTACGAGGAGATTTCTGGCTCTTCAACAATTCAGTGATAACAGCCAGGAGA 60  
 QY 172 ACAACGCTTCTCTAAAGTTCAAGAGGTCCTTTTCCCTGCTGCTGCTGCTGCTGCTGCTG 231  
 Db 61 ACAACGCTTCTCTAAAGTTCAAGAGGTCCTTTTCCCTGCTGCTGCTGCTGCTGCTGCTG 120  
 QY 232 TCTTTGGACTCTAGGAACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 291  
 Db 121 TCTTTGGACTCTAGGAACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 QY 292 GGAATCTTCAGAGATGTGTTTCTGCTGAACCTTGCCTGCTGCTGCTGCTGCTGCTGCTG 351  
 Db 181 GGAATCTTCAGAGATGTGTTTCTGCTGAACCTTGCCTGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 352 CTCTGCCCTTTTGGGCTATGACGACCACTATGATGGGCTTTTGGCAGACAGTCATGTGCA 411  
 Db 241 CTCTGCCCTTTTGGGCTATGACGACCACTATGATGGGCTTTTGGCAGACAGTCATGTGCA 300  
 QY 412 AAATCTTCGAGGATGTATACAACTTCTACGTGTCTCATGCTCACTCTCACTCTGCA 471  
 Db 301 AAATCTTCGAGGATGTATACAACTTCTACGTGTCTCATGCTCACTCTCACTCTGCA 360  
 QY 472 TCACAGTGGATCGTTTCATTTGATGTGTCAGGCTACCAAGGCTTCAACCGCAGGCTA 531  
 Db 361 TCACAGTGGATCGTTTCATTTGATGTGTCAGGCTACCAAGGCTTCAACCGCAGGCTA 420  
 QY 532 AGTGGAGATCTGGGGCCAAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTCTTGTTT 591  
 Db 421 AGTGGAGATCTGGGGCCAAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTCTTGTTT 480  
 QY 592 CTTTGGCCACAGATCATCTATGCGCCATGTTCAAGATATTGAAGCTTATCTGTCATGACC 651  
 Db 481 CTTTGGCCACAGATCATCTATGCGCCATGTTCAAGATATTGAAGCTTATCTGTCATGACC 540  
 QY 652 ACAGTGGAGAGATCTGGGGCCAAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTCTTGTTT 711  
 Db 541 ACAGTGGAGAGATCTGGGGCCAAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTCTTGTTT 600  
 QY 712 CATTGCTCACTATGATCTGCTGCTACT 738  
 Db 601 CATTGCTCACTATGATCTGCTGCTACT 627

RESULT 11  
 CF912382  
 LOCUS  
 DEFINITION A0631A05-5 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)  
 CF912382  
 628 bp mRNA linear EST 05-NOV-2003  
 A0631A05-5 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)  
 cDNA Library (Long 1) Mus musculus cDNA clone NIA:A0631A05

IMAGE:30750532 5', mRNA sequence.  
 CF912382  
 VERSION CF912382.1 GI:38183586  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 628)  
 Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.  
 Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method  
 Genome Res. 11 (9), 1553-1558 (2001)  
 MEDLINE 21429098  
 PUBMED 11544199  
 COMMENT Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Casell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@igsun.grc.nia.nih.gov  
 Plate: A0631 row: A column: 05  
 Seq primer: M13 Reverse  
 High quality sequence stop: 628  
 POLYA=No.

FEATURES  
 source  
 1..628  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CS7BL/6Ncr"  
 /db\_xref="niaEST:A0631A05-5"  
 /db\_xref="taxon:10090"  
 /clone="NIA:A0631A05 IMAGE:30750532"  
 /dev\_stage="Age ~10 weeks old"  
 /lab\_host="DH10B"  
 /clone\_lib="NIA Mouse Hematopoietic Stem Cell (Lin-/C-Kit-/Sca-1+)"  
 /note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Drs. Dennis Taub, Dan Longo (National Institute on Aging, USA), Jonathan Keller (National Cancer Institute, USA). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGAGCGCGCCCTTTTCTTTT-3'] from 1.1 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.2 Kb. The library was constructed by Yulan Piao."

ORIGIN  
 Query Match 50.1%; Score 626.4; DB 14; Length 628;  
 Best Local Similarity 99.8%; Pred. No. 9e-174;  
 Matches 627; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 112 ATGGGCACTACGAGGAGATTCTGGCTTCAACAATTCAGTGAACGCCAGGAGA 171  
 DB 1 ATGGGCACTACGAGGAGATTCTGGCTTCAACAATTCAGTGAACGCCAGGAGA 60  
 QY 172 ACAAAAGCTTCTAAAGTTCAGGAGGCTTTTTCCTGCTGTACCTGGTAGTGTG 231

DB 61 ACAAACGCTTCCTAAAGTTCAGGAGGCTTTTTCCTGCTGTGTACCTGGTAGTGTG 120  
 QY 232 TCTTTGGAGCTGCTAGGAACCTCCCTGCTGCTGATTATATATATATATATATATAT 291  
 DB 121 TCTTTGGAGCTGCTAGGAACCTCCCTGCTGCTGATTATATATATATATATATAT 180  
 QY 292 GGACTCTGACAGATGTGTTTCTGCTGAACCTGGCCCTGGCTGGCTGGTGTGTCTGTA 351  
 DB 181 GGACTCTGACAGATGTGTTTCTGCTGAACCTGGCCCTGGCTGGCTGGTGTGTCTGTA 240  
 QY 352 CTCTGCTCTTTTGGGCTATGCGAGGACCTATGAGTGGTCTTTGGCACAGTCAATGCA 411  
 DB 241 CTCTGCTCTTTTGGGCTATGCGAGGACCTATGAGTGGTCTTTGGCACAGTCAATGCA 300  
 QY 412 AAACCTCTCGAGGCACTGATATACAATGAACCTTACGCTGCTCATGCTCACCTGCA 471  
 DB 301 AAACCTCTCGAGGCACTGATATACAATGAACCTTACGCTGCTCATGCTCACCTGCA 360  
 QY 472 TCACAGTGGATCGTTTCATTGTAGTGTCCAGGCTACCAAGGCTTCAACCGGAGGCTA 531  
 DB 361 TCACAGTGGATCGTTTCATTGTAGTGTCCAGGCTACCAAGGCTTCAACCGGAGGCTA 420  
 QY 532 AGTGAAGATCTGGGGCCAAAGTCAATTTGCTTGTCTCATTTGGTGGTCTCCCTGTTGTTT 591  
 DB 421 AGTGAAGATCTGGGGCCAAAGTCAATTTGCTTGTCTCATTTGGTGGTCTCCCTGTTGTTT 480  
 QY 592 CTTTGGCACAGATCATCTATGGCCATGTTCAAGATATGAAGCTTATCTGTCAGTACC 651  
 DB 481 CTTTGGCACAGATCATCTATGGCCATGTTCAAGATATGAAGCTTATCTGTCAGTACC 540  
 QY 652 ACAGTGAAGATATCCACATGCTGTTCTGTTATACAGATGATCTGGGCTTCTTCTGTC 711  
 DB 541 ACAGTGAAGATATCCACATGCTGTTCTGTTATACAGATGATCTGGGCTTCTTCTGTC 600  
 QY 712 CATTGCTCACTGATGATTCGTGCTACTC 739  
 DB 601 CATTGCTCACTGATGATTCGTGCTACTC 628

RESULT 12  
 CA574451 619 bp mRNA linear EST 19-NOV-2002  
 LOCUS K0616E05-5N NIA Mouse Hematopoietic Stem Cell (Lin-/C-Kit-/Sca-1+)  
 DEFINITION K0616E05-5N NIA Mouse Hematopoietic Stem Cell (Lin-/C-Kit-/Sca-1+)  
 IMAGE:30070324 5', mRNA sequence.  
 CA574451  
 VERSION CA574451.1 GI:25119152  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse).  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 619)  
 Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Luo,A., Carter,M.G., Taub,D., Longo,D.L., Keller,J. and Ko,M.S.H.  
 Systematic Analyses of NIA Mouse Hematopoietic Stem Cell (Lin-/C-Kit-/Sca-1+) cDNA Library (Long)  
 Unpublished (2001)  
 COMMENT Other\_ESTs: K0616E05-3  
 Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Casell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@igsun.grc.nia.nih.gov  
 Plate: K0616 row: E column: 05  
 Seq primer: M13 Reverse  
 High quality sequence stop: 619  
 POLYA=No.

FEATURES  
 source  
 1..619  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CS7BL/6Ncr"

/db\_xref="niaEST:K0616E05-5N"  
/db\_xref="taxon:10090"  
/clone="NIA:K0616E05 IMAGE:30070324"  
/tissue\_type="Hematopoietic Stem Cell  
(Lin-/C-Kit-/Sca-1+)"  
/dev\_stage="Age approx.10 weeks old"  
/lab\_host="DRI0B"  
/clone\_lib="NIA Mouse Hematopoietic Stem Cell  
(Lin-/C-Kit-/Sca-1+) cDNA Library (Long)"  
/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:  
NotI; Mouse cDNA project by the Laboratory of Genetics,  
National Institute on Aging (NIA), Intramural Research  
Program, NIH (http://lsun.grc.nia.nih.gov/cDNA). This is  
a long-transcript enriched cDNA library (Ref. Genome Res.  
11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were  
obtained from Drs. Dennis Taub, Dan Longo (National  
Institute on Aging, USA), Jonathan Keller (National Cancer  
Institute, USA). Double-stranded cDNAs were synthesized  
with an Oligo(dT) primer (Invitrogen).  
5'-pGATAGTCTAGATCGAGCGGCCCTTTTCTTTT-3' from  
1.1 ug of total RNA, treated with T4 DNA polymerase, and  
purified by ethanol-precipitation. The cDNAs were ligated  
to lobe-linker L1-Sal4, purified by phenol/chloroform, and  
separated from free linkers by Centricon 100. Then, the  
cDNAs were amplified by long-range high fidelity PCR using  
Ex Taq polymerase (Takara) with a primer Sal4-S. The  
products were purified by phenol/chloroform and Centricon  
100. The cDNAs were digested with SalI and NotI enzymes  
and cloned into SalI/NotI site of pSPORT1 plasmid vector.  
The DH10B E. coli host was transformed with the ligation  
mixture by the standard chemical method. The average  
insert size is about 2.2 kb. The library was constructed  
by Yulan Piao (NIA)."

Query Match 49.4%; Score 617.4; DB 14; Length 619;  
Best Local Similarity 99.8%; Pred. No. 4.1e-171;  
Matches 618; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 112 ATGGGCACTACAGGAGATTCTGGCTCTTCAACATTCAGTGATACAGCAGGAGA 171  
DB 1 ATGGGCACTACAGGAGATTCTGGCTCTTCAACATTCAGTGATACAGCAGGAGA 60  
QY 172 ACAACAGCTTCTTAAAGTTCAAGAGAGTCTTTTGGCCCTGTGTGTAACCTGGTAGTTTG 231  
DB 61 ACAACAGCTTCTTAAAGTTCAAGAGAGTCTTTTGGCCCTGTGTGTAACCTGGTAGTTTG 120  
QY 232 TCTTTGGACTCTAGGAACCTCCCTGGTCTGTATATATACATTTTCTACAGAGCTGA 291  
DB 121 TCTTTGGACTCTAGGAACCTCCCTGGTCTGTATATATACATTTTCTACAGAGCTGA 180  
QY 292 GGACTCTGACAGATGTGTTTCTGTGAACTTGGCCCTGGCTGACCTGGTGTGTGTTCTGTGA 351  
DB 181 GGACTCTGACAGATGTGTTTCTGTGAACTTGGCCCTGGCTGACCTGGTGTGTGTTCTGTGA 240  
QY 352 CTCTGCCCTTTTGGCCCTATCGAGCACTATAGTGGTCTTTGGCAGCATGTGCA 411  
DB 241 CTCTGCCCTTTTGGCCCTATCGAGCACTATAGTGGTCTTTGGCAGCATGTGCA 300  
QY 412 AAACCTCTCGAGGATGATACATGAACTTCTAGTGTCCATGCTCACTCTCACCTGCA 471  
DB 301 AAACCTCTCGAGGATGATACATGAACTTCTAGTGTCCATGCTCACTCTCACCTGCA 360  
QY 472 TCACAGTGGATCGTTTCAATTTAGTGGTCCAGGTACCAAGCCCTTCAACCGGAGGCTA 531  
DB 361 TCACAGTGGATCGTTTCAATTTAGTGGTCCAGGTACCAAGCCCTTCAACCGGAGGCTA 420  
QY 532 AGTGAAGATCTGGGCGCAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTGTGTTT 591  
DB 421 AGTGAAGATCTGGGCGCAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTGTGTTT 480  
QY 592 CTTTGGCCACAGATCATCTATGGCCATGTTTCAAGATATTGACAAGCTTATCTGTCACTACC 651

ORIGIN

Db 481 CTTTGCCACAGATCATCTATGGCCATGTTCAAGATATTGACAAGCTTATCTGTCTAGTACC 540  
QY 652 ACAGTGAGGAGATATCCACTATGTTCTTTTATACAGATGACTCTGGGTCTCTCTGTC 711  
Db 541 ACAGTGAGGAGATATCCACTATGTTCTTTTATACAGATGACTCTGGGTCTCTCTGTC 600  
QY 712 CATTGCTCACTATGATTCT 730  
Db 601 CATTGCTCACTATGATTCT 619  
RESULT 13  
BB662193  
LOCUS  
DEFINITION  
BB662193 RIKEN full-length enriched, 16 days neonate heart Mus  
musculus cDNA clone D830030F09 5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
BB662193.1 GI:16495948  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 626)  
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,  
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,  
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,  
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,  
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,  
Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.  
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)  
Unpublished (2001)  
TITLE  
JOURNAL  
COMMENT  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp  
URL: http://genome.gsc.riken.go.jp/  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,  
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,  
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.  
and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1759-1771 (2000)  
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,  
Sugahara,Y. and Hayashizaki,Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamana,K., I.,  
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and  
Hayashizaki,Y.  
Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
Please visit our web site (http://genome.gsc.riken.go.jp) for  
further details.  
e mouse tissues.  
Location/Qualifiers  
1. 626  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="D830030F09"

FEATURES  
source

/tissue_type="heart"		/dev_stage="16 days neonate"		/lab_host="DH10B"		/clone_lib="RIKEN full-length enriched, 16 days neonate heart"		/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGCAGAGAGCGCGCAACTCGAGTTTCTTTTNTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGCAGAGATTCTCGAGTTTAAATAATATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."	
ORIGIN		49.2%;	Score 616;	DB 10;	Length 626;				
Query Match		Best Local Similarity	100.0%;	Pred. No. 1.1e-170;	Indels 0;	Gaps 0;			
Matches 61;	Conservative 0;	Mismatches 0;							
QY	15	GGCAGACCTTTAGTGAGCACACTTCACTCTGGAAACAAAGCTACTGGGCTTCTCTTCTGAT	74						
DB	10	GGCAGACCTTTAGTGAGCACACTTCACTCTGGAAACAAAGCTACTGGGCTTCTCTTCTGAT	69						
QY	75	GCCATGATCATGGGCATCAAGTAGTCACTCTGTACGATGGGCACTACGAGGAGGATTC	134						
DB	70	GCCATGATCATGGGCATCAAGTAGTCACTCTGTACGATGGGCACTACGAGGAGGATTC	129						
QY	135	TGCTCTTCAACAATTCAGATGATAACAGCCAGAGAAACAAAGCTTCTCTAAAGTTCAAG	194						
DB	130	TGGCTCTTCAACAATTCAGATGATAACAGCCAGAGAAACAAAGCTTCTCTAAAGTTCAAG	189						
QY	195	GAGGTCCTTTTGGCCCTGTGTACCTCGTAGTGTCTCTTGGAGCTCTAGAAACTCC	254						
DB	190	GAGGTCCTTTTGGCCCTGTGTACCTCGTAGTGTCTCTTGGAGCTCTAGAAACTCC	249						
QY	255	CTGGTCTGATTATATACATTTTCTACCAAGCTGAGGACTCTGACAGATGTGTTTCTG	314						
DB	250	CTGGTCTGATTATATACATTTTCTACCAAGCTGAGGACTCTGACAGATGTGTTTCTG	309						
QY	315	CTGAACCTTGGCCCTGCTGACCTGGTGTGTCTGTACTCTGCGCTTTGGGCGCTATGCA	374						
DB	310	CTGAACCTTGGCCCTGCTGACCTGGTGTGTCTGTACTCTGCGCTTTGGGCGCTATGCA	369						
QY	375	GGCACCCTATGAGTGGGCTTTTGGGCACAGTCATGTGCAAACTCTTTCAGGACATGTATACA	434						
DB	370	GGCACCCTATGAGTGGGCTTTTGGGCACAGTCATGTGCAAACTCTTTCAGGACATGTATACA	429						
QY	435	ATGAACCTTCTACGTGTCCATGCTCACTCTCACCTGCACTACAGTGGATCGTTTCATGTGA	494						
DB	430	ATGAACCTTCTACGTGTCCATGCTCACTCTCACCTGCACTACAGTGGATCGTTTCATGTGA	489						
QY	495	GTGGTCCAGGCTACCAAGGCGCTTCAACCGGACGGCTAAGTGAAGATCTCGGCGCCAAGTC	554						
DB	490	GTGGTCCAGGCTACCAAGGCGCTTCAACCGGACGGCTAAGTGAAGATCTCGGCGCCAAGTC	549						
QY	555	ATTTCGTTGCTCATTTGGGTGGTCTCCCTGTGTGGTTCTTTTGGCACAGATCATCTATGGC	614						
DB	550	ATTTCGTTGCTCATTTGGGTGGTCTCCCTGTGTGGTTCTTTTGGCACAGATCATCTATGGC	609						
QY	615	CATGTTCAAGATATTG	630						
DB	610	CATGTTCAAGATATTG	625						

RESULT 14  
CF910757

```

LOCUS       CF910757                610 bp      mRNA      linear      EST 05-NOV-2000
DEFINITION  A0604D11-5 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)
            CDNA Library (Long 1) Mus musculus cDNA clone NIA:A0604D11
IMAGE:30747982 5', mRNA sequence.

ACCESSION   CF910757
VERSION     CF910757.1  GI:38181933
KEYWORDS    Mus musculus (house mouse)
SOURCE      Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 610)
            Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
            Construction of long-transcript enriched cDNA libraries from
            submicrogram amounts of total RNAs by a universal PCR amplification
            method
            Genome Res. 11 (9), 1553-1558 (2001)
PUBLISHED  21429098
            11544199
COMMENT     Contact: Dawood B. Dudekula
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: A0604 row: D column: 11
            Seq primer: M13 Reverse
            High quality sequence stop: 610
            POLYA=No.

FEATURES             Location/Qualifiers
     1..610
         /organism="Mus musculus"
         /mol_type="mRNA"
         /strain="C57BL/6NCr"
         /db_xref="niaEST:A0604D11-5"
         /db_xref="taxon:10090"
         /clone="NIA:A0604D11 IMAGE:30747982"
         /dev_stage="Age ~10 weeks old"
         /lab_host="DH10b"
         /clone_id="NIA Mouse Hematopoietic Stem Cell
            (Lin-/C-Kit-/Sca-1+) CDNA Library (Long 1)"
         /note="Vector: pCMV-SPORT6 (Invitrogen); Site.1: SalI;
            Site.2: NotI; Mouse cDNA project by the Laboratory of
            Genetics, National Institute on Aging (NIA), Intramural
            Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
            This is a long-transcript enriched cDNA library (Ref.
            Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]. Total
            RNAs were obtained from Drs. Dennis Taub, Dan Longo
            (National Institute on Aging, USA), Jonathan Keller
            (National Cancer Institute, USA). Double-stranded cDNAs
            were synthesized with an Oligo(dt) primer [Invitrogen:
            5'-pGACTAGTCTAGATCGAGCGCGCCCTTTTCTTTT-3'] from
            1.1 ug of total RNA, treated with T4 DNA polymerase, and
            purified by ethanol-precipitation. The cDNAs were ligated
            to Loner-linker IL-SalI, purified by phenol/chloroform, and
            separated from free linkers by Centricon 100. Then, the
            cDNAs were amplified by long-range high fidelity PCR using
            Ex Taq polymerase (Takara) with a primer Sal4-S. The
            products were purified by phenol/chloroform and Centricon
            100. The cDNAs were digested with SalI and NotI enzymes
            and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
            vector. The DH10B E. coli host was transformed with the
            ligation mixture by the standard chemical method. The
            average insert size is about 2.2 kb. The library was
            constructed by Yulan Piao."

```

ORIGIN		constructed by human flag.			
Query Match	48.8%;	Score 610;	DB 14;	Length 610;	
Best Local Similarity	100.0%;	Pred. No. 6.4e-169;			
Matches 610;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	139	TCTTCAACAAATTC	CAGTGATAACAGC	CAGAGAACAAAC	CGTTTCCTTAAAGTTC
					CAAGGAGG 138
Db	1	TCTTCAACAAATTC	CAGTGATAACAGC	CAGAGAACAAAC	CGTTTCCTTAAAGTTC
					CAAGGAGG 60

```

1. .601
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6NCr"
/db_xref="niaEST:A0647A08-5"
/db_xref="taxon:10090"
/clone="NIA:A0647A08 IMAGE:30752071"
/dev_stage="Age ~10 weeks old"
/lab_host="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/C-kit-/Scs-1+) cDNA Library (Long 1)"
/notes="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics' National Institute on Aging (NIA), Intramural
Research Program, NIH (http://igsum.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library [Ref.
Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]]. Total
RNAs were obtained from Drs. Dennis Taub, Dan Longo
(National Institute on Aging, USA), Jonathan Keller
(National Cancer Institute, USA). Double-stranded cDNAs
were synthesized with an Oligo(dT) primer (Invitrogen):
5'-pGACTACTCTCATCGACGGCGGCGCCCTTTTCTTTT-3', and
1.1 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lone-linker LL-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 2.2 kb. The library was
constructed by Yulan Piao."
```

## ORIGIN

Query Match	47.9%	Score 599.4	DB 14	Length 601
Best Local Similarity	99.8%	Mismat. No. 8	7e-166	
Matches 600	Conservative 0	Identical 1	Indels 0	Gaps 0
QY	139	TCTTCAACAATTC	CAGTGATAACAGCCAGGAGAACAAACGGTCTCTAAAGTTC	CAAGGAGG 198
DB	1	TCTTCACAAATTC	CAAGTGATAACAGCCAGGAGAACAAACGGTCTCTAAAGTTC	CAAGGAGG 60
QY	199	TCTTTTGGCCCTG	TGCTGTAACCTGGTAAGTGTGTTGCTTTGGACTGCTAGGAAACTCCCTGG	258
DB	61	TCTTTTGGCCCTG	TGCTGTAACCTGGTAAGTGTGTTGCTTTGGACTGCTAGGAAACTCCCTGG	120
QY	259	TTCTGATTATATACATTTTCTAC	CAGAAGCTGAGGACTCTGCACAGATGTTGTTCTGCTGA	318
DB	121	TTCTGATTATATACATTTTCTAC	CAGAAGCTGAGGACTCTGCACAGATGTTGTTCTGCTGA	180
QY	319	ACTTGCCCTGGCTGACCTGGT	GTGTTGCTGTACTCTGCTGCCCTTTTGGCCCTATGACAGCA	378
DB	181	ACTTGCCCTGGCTGACCTGGT	GTGTTGCTGTACTCTGCTGCCCTTTTGGCCCTATGACAGCA	240
QY	379	CCTATGAGTGGGTCTTTGG	CACAGTCAATGTGCAAAACTCTTCGAGGCATGTATACAATGA	438
DB	241	CCTATGAGTGGGTCTTTGG	CACAGTCAATGTGCAAAACTCTTCGAGGCATGTATACAATGA	300
QY	439	ACTTCTAAGTGTCCANGCTCACTCTCACT	CGCATCACAGTGGATCGTTTCAATTGTAGTGG	498
DB	301	ACTTCTAAGTGTCCANGCTCACTCTCACT	CGCATCACAGTGGATCGTTTCAATTGTAGTGG	360
QY	499	TCCAGGCTTCCAGGCCCTTCAACCG	CAGCGCTTAAGTGGAAAGATCTGGGGCCCAAGTCATTT	558
DB	361	TCCAGGCTTCCAGGCCCTTCAACCG	CAGCGCTTAAGTGGAAAGATCTGGGGCCCAAGTCATTT	420
QY	559	GCTTGCTCAATTGGGNGGCTCCCTGTT	CGTCTTTTGGCCACAGATCAATCTATGGCCATG	618
DB	421	GCTTGCTCAATTGGGNGGCTCCCTGTT	CGTCTTTTGGCCACAGATCAATCTATGGCCATG	480

```

RESULT 15
CF913351
LOCUS
DEFINITION
    CF913351
    A0647A08-5 NIA Mouse Hematopoietic Stem Cell (Lin-/C-Kit-/Sca-1+)
    CDNA Library (long 1) Mus musculus cDNA clone NIA:A0647A08
    IMAGE:30752071 5', mRNA sequence.
    CF913351
    CF913351.1 GI:38184555
    EST.
    Mus musculus (house mouse)
    SOURCE
    CF913351.1
    Mus musculus (house mouse)
    ORGANISM
    Mus musculus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
    P1ao.Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
    1 (bases 1 to 601)
    Construction of long-transcript enriched cDNA libraries from
    submicrogram amounts of total RNAs by a universal PCR amplification
    method
    JOURNAL
    Genome Res. 11 (9), 1553-1558 (2001)
    MEDLINE
    21429098
    PUBMED
    11544199
    COMMENT
    Contact: Dawood B. Dudekula
    Laboratory of Genetics
    National Institute on Aging/National Institutes of Health
    333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
    Email: dcnae@gsun.grc.nia.nih.gov
    Plate: A0647 row: A column: 08
    Seq primer: M13 Reverse
    High quality sequence stop: 601
    POLY=A=No. Location/Qualifiers
FEATURES

```



QY	619	TTCTAGATATTGACAAAGCTTATCTCTCAGTACCCACAGTGGAGATATCCACTATGGTTC	678
Db	481	TTCTAGATATTGACAAAGCTTATCTCTCAGTACCCACAGTGGAGATATCCACTATGGTTC	540
QY	679	TTGTTATACAGATGACTCTGGGGTTCTTCTGCGCCATTGCTCACTATGATTCTGTGCTACT	738
Db	541	TTGTTATACAGATGACTCTGGGGTTCTTCTGCGCCATTGCTCACTATGATTCTGTGCTACT	600
QY	739	C	739
Db	601	C	601

Search completed: September 29, 2004, 02:59:53  
 Job time : 3122 secs

**This Page Blank (uspto)**







```
Db 573 TCTTGGCCACCCAGATGACACTGGGGTCTCTTCTGCCACTGCTGCCAAGATTGTCTGCTA 632
Qy 737 CTCAGGCATTATCAAGACCTTGTCTTCATGCTCGAACTTCCAGAGACACAAATCTCTAAA 796
Db 633 TTCAGTCATAATCAAAACACTGCTTCATGCTGGAGGCTTCCAGAGACACAGATCTCTAAA 692
Qy 797 GATCATCTCTCTAGTGGCTGTCTCTGCTGACCCAGACACCCCTTCAACCTTGGCCAT 856
Db 693 GATCATCTCTCTGTGTGTGTCTCTGCTGACCCAGACACCCCTTCAACCTTGGTAA 752
Qy 857 GTTAATCCAAAGTCAAGCTGGAGTACTATACCATTAACAGAGCTTTAAAGTATGCCATCGT 916
Db 753 GCTCATCGCAGACACACACTGGAGTACTATGCCATGACAGCTTTCACCTACACCATCAT 812
Qy 917 AGTGACAGAGCTATAGCATACTTTCGGGCTTGGCTTAAACCTGTACTTTATGCCCTTGT 976
Db 813 AGTGACAGAGCCATCGCATACCTGAGGGCTGCTTAAACCTGTACTTATGCCCTTGT 872
Qy 977 TGGCTTAAAGTTCGGAAGAACGCTCGAAACTTATGAAGGATATGGCTGCTCTCTCA 1036
Db 873 CAGCTGAAGTTTCAAGAACTTCTGAAACTTGTGAAGACATTTGGCTGTCTCCCTTA 932
Qy 1037 CCTGGAGTCTCAAGTCAATGAAGTCTTCTGAGGACAGTTCOAAGACTTTGTTCTGCCTC 1096
Db 933 CTTTGGGCTCTCACATCAATGGAATCTTCTGAGGACAAATCCAAAGACTTTTCTGCCTC 992
Qy 1097 CCACAACTAGAGACCCAGTATGTTCCAAATCTAG 1133
Db 993 CCACAACTGGAGGCCACAGCATGTTCCAGTTATAG 1029

RESULT 4
US-09-449-437A-1
; Sequence 1, Application US/09449437A
; Patent No. 6319675
; GENERAL INFORMATION:
; APPLICANT: Brisikin, Michael J.
; APPLICANT: Murphy, Kristine E.
; APPLICANT: Wilbanks, Alyson M.
; APPLICANT: Wu, Lijun
; TITLE OF INVENTION: No. 6319675el Antibodies and Ligands for "Bonzo"
; FILE REFERENCE: 1855.1070-000
; CURRENT APPLICATION NUMBER: US/09/449,437A
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-449-437A-1

Query Match 54.3%; Score 679.2; DB 4; Length 1029;
Best Local Similarity 81.2%; Pred. No. 4.1e-215;
Matches 786; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

Qy 166 AGGAGAACAAACGCTTCTTAAGTTCAAGGAGTCTTTTTCGCCCTGTGTACTCTGGTAG 225
Db 62 AGGAGCATCAAGACTTCTCAGTTTCAGCAAGTCTTTTCTGCCCTGCATGTACTCTGGTG 121
Qy 226 TGTTGTCTTTGGACTCTAGGAACCTCCTGTTCTGATTATATATATATCTTACACAG 285
Db 122 TGTTGTCTTTGGTCTGGTGGGAACCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 181
Qy 286 AGCTGAGGACTCTGACAGATGTGTTCTGCTGAACCTTGCCTGGCTGACCTGGTGTGTG 345
Db 182 AGTTGCAGAGCTTGACCGATGTGTTCTGCTGTAACCTTACCCTGGCTGACCTGGTGTG 241
Qy 346 TCTGTACTCTGCCCTTTTGGGCTCTATGACGACCTATGATGGGTCTTTGGACAGTCA 405
Db 242 TCTGTACTCTGCCCTTCTGGGCTCTATGACGACCTATGATGGGTCTTTGGGCGAGTCA 301
```

```
Qy 406 TGTGCAAACTCTTTCGAGGATGTATACAAATGAACCTTCTACGTGTCATGCTACTCTCA 465
Db 302 TGTGCAAAAGAGCTTACTGGGCATCTACACTATTAACTTCTACACGTCTCATCTCACTCA 361
Qy 466 CTTGATCACAAGTGGATCGTTTTCATTGTAGTGGTCCAGGCTTACCAAGGCTTCAACGGGC 525
Db 362 CTTGATCACAAGTGGATCGTTTTCATTGTAGTGGTCCAGGCTTACCAAGGCTTCAACGGGC 421
Qy 526 AGGCTAAGTGAAGATCTGGGGCCAAAGTCAATTCCTGCTCAATTTGGGTGCTTCCCTGT 585
Db 422 AAGCAAGAGGATGACCTGGGGCCAAAGTCAATTCCTGCTCAATTTGGGTGCTTCCCTGT 481
Qy 586 TGGTTTCTTTCGCCACAGATCATCTATGCCATGTTCAGATATTGACAAGCTTATCTGTC 645
Db 482 TGGTTTCTTTCGCCCAAAATTAATCTATGGCAATGCTTTAATCTCGACAGCTCATATGTG 541
Qy 646 AGTACCACAGTGAAGATATCCACTATGTTCTTGTATACAGATGACTTGGGGTCT 705
Db 542 GTTACCATGACGAGGCAATTTCCACTGTGTGTTCTTCCACCAGATGACACTGGGGTCT 601
Qy 706 TCTGCGCAATTCCTACTATGATTCGTGCTACTCAGGATATCAAGACCTTCTGCTCATG 765
Db 602 TCTTGGCACTGCTCACCATGATTTGTGCTATTGCTGCTATTCAATCAAAACACTGCTTTCATG 661
Qy 766 CTCGAACTTCCAGAGCACAAATCTCTAAAGATCATCTTCTTGTAGTGGCTGTGTTC 825
Db 662 CTGGAGGCTTCCAGAGCACAGATCTCTAAAGATCATCTTCTGCTGATGCTGTGTTC 721
Qy 826 TGTGACCCAGACACCTTCAACCTTGGCCATGTTTAACTCAAGTACAGCTGGGAGTACT 885
Db 722 TGTGACCCAGATGCCCTTCAACCTCATGAAGTTTCACTCCGAGCACACACTGGGAATACT 781
Qy 886 ATACCATTAACAGCTTTAAGTATGCCATCGTAGTACAGAGGCTATAGCATCTTCCGG 945
Db 782 ATGCCATGACAGCTTTCATCTACCATCATGTTGACAGAGGCTATGCATACCTGAGG 841
Qy 946 CTTGCTTAAACCTGTACTTTTATGCTTTGTGCTTAAAGTTCCGGAAGACCTCTGGA 1005
Db 842 CTTGCTTAAACCTGTGCTCTATGCTTGTGCTTAAAGTTTCAAGTTTTCGAAAGAACTTCTGGA 901
Qy 1006 AACCTATCAAGGATATGGCTGCTCTCTACCTGGGAGTCTCAAGTCAATGGAAGCTT 1065
Db 902 AACCTTGAAGAGATGTTGTTGCTTCCCTTACCTTGGGCTCTCATCAATGGAATCTT 961
Qy 1066 CTGAGGACAGTTCGAAGACTTGTCTGCTTCCCAATGTAGAGACCAAGTATGTTC 1125
Db 962 CTGAGGACAAATTCGAAGACTTGTCTGCTTCCCAATGTAGAGACCAAGTATGTTC 1021
Qy 1126 AATTGTAG 1133
Db 1022 AGTTATAG 1029

RESULT 5
US-09-275-384B-4
; Sequence 4, Application US/09275384B
; Patent No. 6232084
; GENERAL INFORMATION:
; APPLICANT: MACPHER, COLIN HOUSTON
; APPLICANT: MOORES, KITTY
; TITLE OF INVENTION: NEW USE
; FILE REFERENCE: GH-31106
; CURRENT APPLICATION NUMBER: US/09/275,384B
; CURRENT FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: 980677.2
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
```

US-09-275-384B-4

Query Match 54.1%; Score 676.8; DB 3; Length 1026;  
Best Local Similarity 81.2%; Pred. No. 2.5e-214;  
Matches 783; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 166 AGAGAACAAACGCTTCTTAAAGTTCAGAGAGTCTTTTGGCCCTGTGTGACTCTGGTAG 225  
DB 62 AGAGCATCAAGACTTCTTGCAGTTCAGCAAGTCTTCTGCCCTGCATGTACCTGTGG 121

QY 226 TGTGTGCTTTGGACTGTGAGAACTCCTCGTCTGATATATACATTTTCTACACAGA 285  
DB 122 TGTGTGCTGTGTGTGGTGGGAACTCTGTGTGCTGTATATCAATCTTCTACATA 181

QY 286 AGCTAGGACTCTGACAGATGTGTTTCTGCTGAATTCGCCCTGCTGACCTGTGTTG 345  
DB 182 AGTTCAGAGCCTGACGAGTGTCTTCTGGTGAACCTACCCCTGGCTGACCTGTGTTG 241

QY 346 TCTGTACTCTGCCCTTTTGGGCTATGAGGACCTATGAGTGGTCTTTGGCAGTCA 405  
DB 242 TCTGCACCTCTGCCCTTCTGGGCTATGAGGACCTATGAGTGGTCTTTGGCAGGTC 301

QY 406 TGTGCAAACTCTTGGAGCATGTATACAATGAATTTCTAGTGTCCATGCTCACTCTCA 465  
DB 302 TGTGCAAGAGCCTACTGGGCATCTACACTATTAACTTCTACAGTCCATGCTCATCTCA 361

QY 466 CTGTGATACAGTGTGATGTTTCAATGTAGTGTGCTGAGCTTACCAAGCCTTCAACCGC 525  
DB 362 CTGTGATCACTGTGATGTTTCAATGTAGTGTGTTAAGGCCACCAAGCCTTCAACCCAG 421

QY 526 AGCTAAGTGAAGATCTGGGGCCAACTTCTTGTCTCATTTGGTGTCTCCCTGT 585  
DB 422 AGCCAAAGAGTACCTGGGGCAAGTCAACAGCTTCTCATCTGGTGTATCCCTGC 481

QY 586 TGGTTCTTTGTCACAGATCATCTATGGCAATGTTCAAGATTTGACAAGCTTATCTGTC 645  
DB 482 TGGTTCTTTGTCACAAATATCTATGGCAATGTTCTTAACTTCGACAAGCTATATGTG 541

QY 646 AGTACACAGTGAAGATATCCATGTTGTTTATAGATGATCTGAGTCTGGGTTCT 705  
DB 542 GTTACCATGACAGGCAATTTCCATGTGTTTCTGCCCACCAAGATGACATGGGGTCT 601

QY 706 TCTGCCATTTGCTACTATGATTTCTGTCTACTCAGGCAATATCAAGACCTTGTCTCATG 765  
DB 602 TCTGCCATTTGCTACCATGATTTCTGTATTCAGTCAATCAAAACACTGTCTTCATG 661

QY 766 CTCGAACTTCAGAACACAAATCTTAAAGATCATCTTCTGTAGTGTGCTGTGTTCC 825  
DB 662 CTGGAGGCTTCAGAAAGCACAGATCTTAAAGATCATCTTCTGTGTGATGCTGTGTTCC 721

QY 826 TGTGACCCAGACACCTTCAACCTTGGCATGTTTAAATCAAGTACAAGCTGGAGTACT 885  
DB 722 TGTGACCCAGATGCTTCAACCTCATGAGTTCTTCGAGACACACACTGGGATACT 781

QY 886 ATACCAATCAACGCTTTAAGTATGCTAGTGAAGAGCTTATAGCATACTTTCGGG 945  
DB 782 ATGCCATGACCACTTCTACTACACCATCATGTTGACAGAGGCCATCATACCTGAGG 841

QY 946 CTGTGCTTAAACCTGCTACTTATGCTTGTGCTTAAAGTTCGGAAGACCTCTGGA 1005  
DB 842 CTGTGCTTAAACCTGCTCTATGCTTGTGCTGCTGAAAGTTTCGAAAGAACTTCTGGA 901

QY 1006 AACTTATGAAGATATCGGCTGCTCTCTCACTCGGAGTCTCAAGTCAATGGAAGTCT 1065  
DB 902 AACTTGTGAAGACATTTGTTGCCCTCCCTTACCTTGGGGTCTCATCAATGGAATCT 961

QY 1066 CTGAGACAGTTCAGAGCTTGTCTGCTCCCAACAATGAGAGACCAAGTATGTTCC 1125  
DB 962 CTGAGACAAATTCAGAGCTTTTCTGCTCCCAACAATGAGAGACCAAGTATGTTCC 1021

QY 1126 AATT 1129  
DB 1022 AGTT 1025

RESULT 6  
US-09-116-498-1  
; Sequence 1, Application US/09116498  
; Patent No. 6251582  
; GENERAL INFORMATION:  
; APPLICANT: Littman, Dan R.  
; Deng, Hongkui  
; Unutmaz, Derya  
; Ramani, Vineet N.K.  
; TITLE OF INVENTION: NOVEL ALTERNATIVE G-COUPLED RECEPTORS  
; ASSOCIATED WITH RETROVIRAL ENTRY INTO CELLS, METHODS OF  
; IDENTIFYING THE SAME, AND DIAGNOSTIC AND THERAPEUTIC USES  
; THEREOF  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/116,498  
; FILING DATE: 16-Jul-1998  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1049-1-009  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1026 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-116-498-1

Query Match 51.8%; Score 648; DB 3; Length 1026;  
Best Local Similarity 80.8%; Pred. No. 9.4e-205;  
Matches 805; Conservative 0; Mismatches 185; Indels 6; Gaps 4;

QY 141 TTCAACAATTCACAGTGTATACAGCC--AGGAGAACAAACGCTTCTTAAAGTTCAAGGAG 197  
DB 34 TTACAGAGTTTCAATGACACAGCAGGAGGAGCATCAAGACTTCTGCAAGTTCAAGAG 93

QY 198 GTCTTTTGGCCCTGTGTGACTCTGGTAGTGTGTTCTTTGGAGTGTAGGAAACTCCCTG 257  
DB 94 GTCTTTTGGCCCTGTGTGACTCTGGTAGTGTGTTCTTTGGTGTGGTGGGAACTCTCTG 153

QY 258 GTTCTGATATATACATTTTCTACAGAGCTGAGGACTCTGACAGATGTTTCTGCTG 317  
DB 154 G-TGTGCTCATATCCATCTTCTACCAATAAGTTCCAGAGCCTGACGATGTTTCTCTG 211

QY 318 AACTTGGCCCTGCTGACCTGTGTTGTGTGCTGCTGCTCTGCTGCTGCTGCTGCTGCTG 377  
DB 212 AACTTACCCCTGCTGACCTGTGTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 271

378 ACCTATGAGTGGGCTTTGGACAGTCATGTCGAAACTCTTCAGGCATGTATCAATG 437  
 Db ATCCATGAATGGGTGTTGGCCAGGTCTATGTGCAAGAGCCCTAGG-GGCATCTACATAT 330  
 438 AACTTCTACGTGTCATGCTCACTCTCACTGATGATGATGATGATGATGATGATG 497  
 Db AACTTCTACGTGTCATGCTCACTCTCACTGATGATGATGATGATGATGATGATG 390  
 498 GTCCAGGCTACCAAGGCTTCAACCGGAGGCTAAGTGAAGATCTGGGCGCAAGTCAIT 557  
 Db GTTAAGGCCCAAGGCTTCAACCGGAGGCTAAGTGAAGATCTGGGCGCAAGTCAIT 450  
 558 TGCTTGCTCATTTGGGTGGTCTCCCTGTTGGTTCCTTTCGACAGATCATCTATGCCC 617  
 Db AGCTTGCTCATCTGGGTGATATCCCTGCTGTTTCTTGGCCCAAAATATCTATGCAAT 510  
 618 GTTCAAGATATGCAAGCTTATCTGTCAGTACACAGTACGAGAGATATCCATATGGT 677  
 Db GTCTTTAATCTCGACAAGCTCATATGTTTACCATGACGAGGCAATTTCCACTGTGGT 570  
 678 CTTGTTATACAGATGATCTGGGGTCTTCTGCGCATGCTCACTATGATCTGTGCTAC 737  
 Db CTTGCCACCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630  
 738 TCAGGCAATATCAAGACCTTCTTCTGCTGCAAACTTCCAGAGGACCAATCTCTAAG 797  
 Db TCAGTCAATATCAAACTGCTTCTGCTGCAAACTTCCAGAGGACCAATCTCTAAG 690  
 798 ATCATCTTCTGTTAGTGGTCTTCTGCTGCAAACTTCCAGAGGACCAATCTCTAAG 857  
 Db ATCATCTTCTGTTAGTGGTCTTCTGCTGCAAACTTCCAGAGGACCAATCTCTAAG 750  
 858 TTAATCCAAAGTACAAGCTGGGAGTACTATACCATACCAAGCTTAAAGTATGCCATG 917  
 Db TTAATCCAAAGTACAAGCTGGGAGTACTATACCATACCAAGCTTAAAGTATGCCATG 810  
 918 GTGACAGAGGCTATAGCATATTTGGGGCTTGCCTTAACCTGCTGCTTATGCTTTGTT 977  
 Db GTGACAGAGGCTATAGCATATTTGGGGCTTGCCTTAACCTGCTGCTTATGCTTTGTT 870  
 978 GGCTTAAAGTTCGGAGAGAGCTGCGAACTATGAGAGATATGAGAGATATGAGAG 1037  
 Db AGCCTGAAGTTTCGAAAGAACTCTGGAAGAACTTGTGAAGAGACATTTGGTCCCTTAC 930  
 1038 CTGGGAGTCTCAAGTCAATGAACTTCTGAGGACAGTTTCAAGACTTGTCTGCTCC 1097  
 Db CTGGGAGTCTCAAGTCAATGAACTTCTGAGGACAGTTTCAAGACTTGTCTGCTCC 930  
 1098 CACAATGTAGAGACACCAAGTATTTCCAAATGTAG 1133  
 Db CACAATGTAGAGACACCAAGTATTTCCAAATGTAG 1026

RESULT 7  
 US-09-170-496D-23  
 ; Sequence 23, Application US/09170496D  
 ; Patent No. 655339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Behan, Dominic P.  
 ; APPLICANT: Chalmers, Derek T.  
 ; APPLICANT: Liaw, Chen W.  
 ; TITLE OF INVENTION: No. 655339-Endogenous, Constitutively Activated Human G Protein-  
 ; FILE REFERENCE: AREN-0040  
 ; CURRENT FILING DATE: 1998-10-13  
 ; NUMBER OF SEQ ID NOS: 294  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 23  
 ; LENGTH: 1074  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-170-496D-23

Query Match 14.7%; Score 183.4; DB 4; Length 1074;  
 Best Local Similarity 52.9%; Pred. No. 2.3e-50;  
 Matches 490; Conservative 0; Mismatches 401; Indels 36; Gaps 3;  
 146 CAATTCAGTGTATCAAGCCAGGAGACAAACGCTTCTTAAAGTTCAAGGAGTCTTTT 205  
 Db CAATTCAGTGTATCAAGCCAGGAGACAAACGCTTCTTAAAGTTCAAGGAGTCTTTT 116  
 206 GCCCTGTGTGACCTGTGAGTGTGTTTGTCTTGTGACTGTCTAGGAACTCCCTGTGTTGAT 265  
 Db GCCCTGTGTGACCTGTGAGTGTGTTTGTCTTGTGACTGTCTAGGAACTCCCTGTGTTGAT 176  
 266 TATATACATTTTACCAAGAGCTGAGGACCTCTACAGATGTGTTTCTGTGACTGTGCT 325  
 Db TATATACATTTTACCAAGAGCTGAGGACCTCTACAGATGTGTTTCTGTGACTGTGCT 236  
 326 CTTGCTGACCTGTGTTTGTCTGTACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 385  
 Db CTTGCTGACCTGTGTTTGTCTGTACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 296  
 386 GTGGTCTTTTGGACAGTCTGTCGAAACTCTTTCGAGGATGTATACATGACTTCTA 445  
 Db GTGGTCTTTTGGACAGTCTGTCGAAACTCTTTCGAGGATGTATACATGACTTCTA 356  
 446 CGTGTCCATGCTCACTCTCACTCTGATGATGATGATGATGATGATGATGATGATG 505  
 Db CGTGTCCATGCTCACTCTCACTCTGATGATGATGATGATGATGATGATGATGATG 416  
 506 TACCAAGGCTTTCAACCGGAGGCTTAAAGTGAAGATCTGGGGCCCAAGTCTATTTGCT 565  
 Db TACCAAGGCTTTCAACCGGAGGCTTAAAGTGAAGATCTGGGGCCCAAGTCTATTTGCT 476  
 566 CATTTGGGTGCTCTCCCTGTTGTTTCTTCCACAGATCATCTATGCGCATGTTCAAGA 625  
 Db CATTTGGGTGCTCTCCCTGTTGTTTCTTCCACAGATCATCTATGCGCATGTTCAAGA 536  
 626 T-----ATTGCAAGCTTATCTGTCACTCAACAGTGGAGATATCCACTATG-- 674  
 Db T-----ATTGCAAGCTTATCTGTCACTCAACAGTGGAGATATCCACTATG-- 596  
 675 -----GTTCTGTATACAGATGATCTCTGGGTTCTTCTCTGCTTCTCTCTCTCTCT 727  
 Db -----GTTCTGTATACAGATGATCTCTGGGTTCTTCTCTGCTTCTCTCTCTCTCT 656  
 728 TCTGTGCTTACTCAGGCAATTTCAAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 787  
 Db TCTGTGCTTACTCAGGCAATTTCAAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 716  
 788 ATCTCTAAGATCATCT 847  
 Db ATCTCTAAGATCATCT 776  
 790 TGTCTGCTTATACCATCATCTTCAACCTGATACAGCCCAAGAGTCTTCCAAAGCACA 907  
 Db TGTCTGCTTATACCATCATCTTCAACCTGATACAGCCCAAGAGTCTTCCAAAGCACA 836  
 908 TGCCA-----TCGTAGTGACAGGCTATAGCATACTTCTCTCTCTCTCTCTCTCTCT 949  
 Db TGCCA-----TCGTAGTGACAGGCTATAGCATACTTCTCTCTCTCTCTCTCTCTCT 896  
 950 CTTTAACTCTGTACTTATGCTTGTGCTTAAAGTTCGGAAGAACGCTCTGGAAC 1009  
 Db CTTTAACTCTGTACTTATGCTTGTGCTTAAAGTTCGGAAGAACGCTCTGGAAC 956  
 1010 TATGAAGATATCGGCT 1036  
 Db TATGAAGATATCGGCT 983

RESULT 8  
 US-09-266-464-1  
 ; Sequence 1, Application US/09266464



GENERAL INFORMATION:  
APPLICANT: Andrew, David P.  
APPLICANT: Zabel, Brian A.  
APPLICANT: Penath, Paul D.  
TITLE OF INVENTION: ANTI-GPR-9-6 ANTIBODIES AND METHODS OF IDENTIFYING AGENTS WHICH MODULATE GPR-9-6 FUNCTION  
FILE REFERENCE: LXS98-16  
CURRENT APPLICATION NUMBER: US/09/266,464  
CURRENT FILING DATE: 1999-03-11  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 2577  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (58)...(1131)  
US-09-266-464-1

Query Match 14.7%; Score 183.4; DB 4; Length 2577;  
Best Local Similarity 52.9%; Pred. No. 4.1e-50;  
Matches 490; Conservative 0; Mismatches 401; Indels 36; Gaps 3;  
QY 146 CAATTCACGCTGATACAGCCAGAGAGAAACAAACGCTTCCTTAAAGTTCAAGGAGGTCTTTT 205  
DB 114 CAACTTCAGTCTACTCTGTGAGAAACAAATGTCCAGGAGTTGGAGCCATTTCT 173  
QY 206 GCCGTGTGTACCTGGTAGTGTGTTCTTTGGAGCTGTAGGAACTCCCTGGTTCTGAT 265  
DB 174 CCACCCCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 233  
QY 266 TATATACATTTCTACCAAGCTCAGGACTGACAGATGTTCTGCTGAATGCC 325  
DB 234 TGTCTACTGTTACTGCAAGAGTAGAACCAACAGACATGTTCTTTTGAATTTGGC 293  
QY 326 CTGGCTGACCTGGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 385  
DB 294 AATTGCTGACCTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 353  
QY 386 GTGGCTGTTGGCAGATGATGCAAACTCTTCGAGCATGTATACAAATGACTTCTA 445  
DB 354 GTGGAAGTTCAGACCTTCACTGCAAGGTGCTCAACAGCATGTACAAAGATGAACTCT 413  
QY 446 CGTGTCCATGCTCACTCTCACTGCTGATCAGATGATGTTCTGTTAGTGTGCTCAGGC 505  
DB 414 CAGCTGTGTTGCTGATCATGTCATCAGCTGACAGGTACATGCTGCTGCTGCTGCTG 473  
QY 506 TACCAAGGCTTCAACCGGAGCTTAAGTGAAGATCTGGGGCCAAAGTCATTTGCTGCT 565  
DB 474 CATGAGACACATCTTGGAGGGAGAAAGGCTTTGTACAGCAAAATGGTTGCTTTAC 533  
QY 566 CATTTGGGTGCTCCCTGTTGTTCTTGTCCACAGATCATCTATGCGCATGTTCAAGA 625  
DB 534 CATCTGGTATGGGAGCTGCTCTGCTGATCCAGAAATCTTATACGCCAAATCAAGA 593  
QY 626 T-----ATTGCAAGCTTATCTGTGCTGATCAGTACCAAGTGTGAGGAGATATCACTATG-- 674  
DB 594 GGAATCCGGCATTTGCTATCTGACCATGTTTACCTAGCGATGAGACCAAACTGAA 653  
QY 675 -----GTTCTGTTATACAGATCACTCTGGGTCTCTGCTGCTGCTGCTGCTGCTGCT 727  
DB 654 GTCAGTGTCTGACCCCTGAAAGGTCACTTCTGGGGTCTCTCTCTCTCTCTCTCTCTCT 713  
QY 728 TCTGTGCTACTCAGGCATTATCAGACCTTGTCTCATGCTCGAACTTCCAGAGACCAA 787  
DB 714 TTGCTGCTATACCATCATCATTCATCACCCTGATACAGCCAGAGTCTTCCAGACAA 773  
QY 788 ATCTCTAAAGATCATCTCTGTTAGTGGCTGTGTTCTGCTGAGCCAGACCTCTTCAA 847  
DB 774 AGCCCTAAAGTGACCATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 833  
QY 848 CCTTGCCATGTTAATCCAAAGTACAGCTGGGAGTACTATACCATACACGCTTTAAGTA 907

DB 834 CTGCATTTTGTGTGTCAGACCATTTGACGCTATGCCATGTTCACTCCCAACTGTGCGGT 893  
QY 908 TGCCA-----TCGTAGTGACAGAGGCTATAGCATACTTTGGGCTTG 949  
DB 894 TTCACCAACATTCACATCTGCTTCAGGTCCACAGACCATGCTCTTCTTCCACAGTTG 953  
QY 950 CTTTAAACCTGTACTTTTATGCTTTGCTTTAAAGTTCCGGAAGAACCTCTGGAACCT 1009  
DB 954 CCTGAACCTGTTCTCTATGTTTGTGGTGAGAGATTCGCGCGGATCTCGTGAAC 1013  
QY 1010 TATGAAGGATATCGCTGCTCTCTCA 1036  
DB 1014 CCTGAAGAACTTGGTTCATCAGCCA 1040  
RESULT 9  
US-09-016-434-1095  
Sequence 1095, Application US/09016434  
Patent No. 6500938  
GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,434  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1095:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2577 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: G1245054  
US-09-016-434-1095

Query Match 14.7%; Score 183.4; DB 4; Length 2577;  
Best Local Similarity 52.9%; Pred. No. 4.1e-50;  
Matches 490; Conservative 0; Mismatches 401; Indels 36; Gaps 3;  
QY 146 CAATTCAGTGTATACAGCCAGAGAGAAACAAACGCTTCCTTAAAGTTCAAGGAGGTCTTTT 205  
DB 114 CAACTTCAGTCTACTCTGTGAGAAACAAATGTCCAGGAGTTGGAGCCATTTCT 173  
QY 206 GCCCTGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 265

Db	174	CCCA	CCCTTGTACTGGCTCGTGTTCATCGTGGTGCCCTTGGGCACACAGTCTGTGTATCCT	233
Qy	266	TATATA	CATTTCCTPACCAGAAAGCTGAGGACTCTGACAGATGTGTTTCGTGTAACCTTGGC	325
Db	234	TGTCTA	CTGCTACAGAGTGAAGACCATGACGCACATGTCCTTTTGATATTGGC	293
Qy	326	CCTGCTGA	CCCTGGTGTTCCTGTACTCTGCGCTTTTGGGCTTATGACGACCACTTAAGA	385
Db	294	AATTGCTG	ACCTCTCTTCTGTCACTCTTCCCTCTGGCCATTGCTGCTGTGACCA	353
Qy	386	GTGGGTCT	TTTGGCACAGCTCATGTGCAAACTCTTCGAGGCATGATATAAATGAACCTCTA	445
Db	354	GTGGAAGTT	COAGACCTTCATGTGCAGAGGTGTCAACAGCATGTACAGATGAACCTCTA	413
Qy	446	CGTGCCATG	CTCACTCTCACTGCAATCAAGTGGATCGTTTCATGTAGTGTCCAGGC	505
Db	414	CAGCTGTGT	GTGCTGATCATGTGCATCAGCGGTGACAGGTACATTCGCATTGGCCAGGC	473
Qy	506	TACCAAGCC	TTCAACCGGCAGCTAAGTGGAGATCTGGGSCCAAGTCATTTCCTCTTGCT	565
Db	474	CATGAGAC	CACTACTTGGAGGAGAAAGCTTTGTACAGCAAAATGGTTTCGTTTAC	533
Qy	566	CATTGGGTG	GTCTCCCTGTGTGTTCTTTTGGCACAGATCATCTATAGGCCCATGTTCAGA	625
Db	534	CATCTGGGT	ATTGGACGTCTCTCTGCATCCAGAAATCTTATACAGCCAAATCAAGGA	593
Qy	626	T-----	ATTGCACAGCTTATCTGCAGTACCACAGTGAAGAGATATCCACTATG--	674
Db	594	GGAA	TCGGGCAATCTATCTGACCATGGTTTACCCTAGCGATGAGACCAACCTGAA	653
Qy	675	-----	GTTCTTGTATACAGATGACTCTGGGGTTCCTTCGTGCATTCCTCACTATGAT	727
Db	654	GTACGTGT	CTTGACCTGAAGTCAATCTGGGGTTCCTTCCTTCCTTCGTGGTCAATGGC	713
Qy	728	TCTGTCTACT	CAGGCATATCAAGACCTTGCTTCATGCTCGAACTTCCAGAGACACAA	787
Db	714	TGTGCTCTA	TACCATCATCATTCACACCTGATACAGCCAAAGAGTCTTCCAGACACAA	773
Qy	788	ATCTCTAA	GATCATCTTCCTTGTAGTGGCTGTGTTCTGTGACCCACAGACACCTTCAA	847
Db	774	AGCCCTAA	AGTGACCATCACTGCTGACCGTCTTGTGCTGTCTCAGTTTCCTCTACAA	833
Qy	848	CTTGCCCATG	TATCCAAAGTACAGCTGGGAGTACTATACCATAAACAGCTTTTAAGTA	907
Db	834	CTGCATTT	TGTGTGTCAGACCATTCAGCGCTATGCGCTTCACTCTCAACTGTGCGGT	893
Qy	908	TGCCA-----	TCGTAGTGACAGAGGCTATAGCATACTTTCGGGGTTG	949
Db	894	ITCCACCA	CAACTTGCACTGCTTCCTCAGGTCAACCCAGACCATCGCTTCTTCCACAGTTG	953
Qy	950	CCTTAAC	CCCTGACTTATTCGCTTCTGTTAAGTTTCGGAGAGACGCTCGGAACT	1009
Db	954	CCTGAAC	CCCTGTTCTCTATGTTTTGTGGGTGAGAGATTCCCGCGGGATCTCTGTGAAAC	1013
Qy	1010	TATGAAG	GATATCGGCTGCCTCTCTCA	1036
Db	1014	CCTGAAGA	ACTTGGTTGCAATCAGCCA	1040

RESULT 10  
US-08-153-848-23  
; Sequence 23, Application US/08153848  
; Patent No. 5759804  
; GENERAL INFORMATION:  
; APPLICANT: Godiska, Ronald  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Schweikart, Vicki L.  
; TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell

STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/153,848  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/977,452  
FILING DATE: 17-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5759804and, Gresta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31794  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2751 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: intron  
LOCATION: 1..691  
FEATURE:  
NAME/KEY: exon  
LOCATION: 692..1771  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 692..1768  
FEATURE:  
NAME/KEY: polyA signal  
LOCATION: 2341..2348  
US-08-153-848-23

Query Match	14.5%	Score 181.2	DB 1	Length 2751
Best local Similarity	53.2%	Pred. No. 2.3e-49		
Mismatches	0	Mismatches 383	Indels 40	Gaps 3
171	AACAAAGCCTCTCTAAAGTTCAAGAGGCTCTTTTGGCCCTGTGTGTCACCTGGTAGTCTTT	230		
779	AAGAAGGATGTGGGAACCTTAAAGCCTGGTTCCTGCCTCATGTATCTGTCAATCGC	838		
231	GTCTTTGGCATGCTAGGAAACTCCCTGGTTCGTATTATATACATATTTCTACCAAGACTG	290		
839	TTCTGTGGCCTGCTCGGCAACGGCTGGTGATCTAGCATACATCTATTTCAGAGGCTC	898		
291	AGGACTCTGACAGATGTGTTTCTGTGTAATTCGCCCTGGCTGGCTACCTGGTGTGTTGTCGT	350		
899	AAGACCATGACGGATACCTACCTGCTCAACCTGGCCGTGGCAGACATCCTTTTCTCTCTA	958		
351	ACTCTGCCCTTTTGGGCCCTATGACGGCACTATGAGTGGTCTTTGGCACAGTCAATGTC	410		
959	ATTCTTCCCTCTGGGCCCTACAGCAAGCAAGTCTCGATCTTTGGCGTCTACTCTGT	1018		
411	AAACTCTTCAGGCGATGTATCAATGAACCTTCTACGTGTCCATGCTCACTCTCAACCTGC	470		
1019	AAGGCACTTTTGGCATCTAATAGTTAAGCTTCTTCAGGGGATGCTGCTGCTCTATGC	1078		
471	ATCACAGTGGATGTTTTCATTGTAGTGGTCCAGGCTACCAAGGCCTTCAACCGCAGGCT	530		
1079	ATCGCATTCACCTGCTACCTAGCATCTGCTCCAGGCCGTGTGGCGTCATCGCACCGCGCT	1138		

```
QY 531 AAGTGAAGATCTGGGGCCAACTCATTTGCTTGTCTCATTTGGTGGTCTCCTCTGTTGTT 590
Db 1139 CCGTGTCTCTCATCAGCAAGCTGTCTGTGGGCACTCTGGATCTGGCCCTCTTCTC 1198
QY 591 TCTTTGCCACAGATCATCTATGCCATGTTCAAGATA-----TTGACA 633
Db 1199 TCATCCCGGAGCTGCTCTACAGCGCTCCAGAGACAGCGGGGAGGACAGCTGAGA 1258
QY 634 AGCTTATCTGTCAGTACACAGTGAAGAGATATCCACTATGTTCTTTTATACAGATGA 693
Db 1259 TGCTCACTGGTCACTGCCCAAGT--GGAGGCGCTTGATCACCATCCAAAGTGGCCAGATGG 1316
QY 694 CTCTGGGTTCTCTGCCATTGCTCACTATGATTTCTGTCTACTCAGGCATATTCAGA 753
Db 1317 TTTTGGGTTCTAGTGCTATGCTGGCTATGAGTTCTGTACTTATTCATCCGTA 1376
QY 754 CTTGCTTTCATCTCGAACTTCCAGAGACAAATCTCTAAAGATCATCTTCTCTGTAG 813
Db 1377 CTTGCTCAGGACGCAACTTTGAGCGGAACAAGGCCATCAAGGTGATCATTCGCGTGG 1436
QY 814 TGGCTGTGTTCTGTGACCGAGACACCTTCAACC-----TTG 852
Db 1437 TGGTAGTCTTCATAGTCTTCAGCTGCGCTCAATGCGGCTGCTGCTCAGACGGTGG 1496
QY 853 CCATGTTAATCAAAAGTACAAGCTGGAGTACTATACCATTAACCAAGCTTAAAGTATGCA 912
Db 1497 CCAACTTCAACATCACCATTAGCAGCTGTGCGAACCAGCAAGCAGCTCAACATTGCT 1556
QY 913 TCGTAGTGACAGAGCTATAGATCATCTTTGGGGCTTGGCTTAACTCTATCTTTATGCT 972
Db 1557 ATGAGCTCACCTACAGCTGGCTCGCTCGCTGCTGCTGCTCAACCTTTCTTTGATGCT 1616
QY 973 TTGTTGGCTTAAAGTTCGGAAGAGAGCTGCGAACTTATGAAGATATCGGCTGCTCT 1032
Db 1617 TCATCGGCTCAAGTTCGCGACGACCTTCTCAAGCTCTTCAAGACTTTGGGCTGCC 1676
QY 1033 CTC 1036
Db 1677 GCA 1680
```

## RESULT 11

```
US-09-299-843A-23
; Sequence 23, Application US/09299843A
; Patent No. 6107475
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweikart, Vicki L.
; TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
```

```
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; City: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
```

## COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,843A
```

## FILING DATE:

```
; CLASSIFICATION: 435
```

## PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: US 09/088,337
; FILING DATE: 01-JUN-1998
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 08/153,848
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,452
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Jill E. Uhl
; REGISTRATION NUMBER: 43,213
; REFERENCE/DOCKET NUMBER: 27866/32059B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX:
```

## INFORMATION FOR SEQ ID NO: 23:

## SEQUENCE CHARACTERISTICS:

```
; LENGTH: 2751 base pairs
```

```
; TYPE: nucleic acid
```

```
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: DNA (genomic)
```

## FEATURE:

```
; NAME/KEY: intron
```

```
; LOCATION: 1..691
```

## FEATURE:

```
; NAME/KEY: exon
```

```
; LOCATION: 692..1771
```

## FEATURE:

```
; NAME/KEY: CDS
```

```
; LOCATION: 692..1768
```

## FEATURE:

```
; NAME/KEY: polyA signal
```

```
; LOCATION: 2341..2348
```

```
US-09-299-843A-23
```

## Query Match

```
14.5%; Score 181.2; DB 3; Length 2751;
```

```
Best Local Similarity 53.2%; Pred. No. 2.3e-49;
```

```
Matches 481; Conservative 0; Mismatches 383; Indels 40; Gaps 3;
```

```
QY 171 AACAAACGCTTCCTAAAGTTCAAGGAGGCTTTTTTGGCCCTGTGTACCTGGTAGTGT 230
Db 779 AAGAGGATGTGGGAACITTAAGGCTGTCTCTGCTCATGTATTCTGTCATGCG 838
QY 231 GTCTTTGGACTGTAGAAACTCCCTGGTTCTGATTATATATACATTTTCTACCAAGAGCTG 290
Db 839 TTCGTGGGCTGTCTCGGCAACGGGCTGTGATACCTGACGTACATCTATTTCAGAGGCTC 898
QY 291 AGGACTCTGACAGATGTGTTTCTGTGCACTTCCCTCGCTGACCTGGTGGTGTGCTGT 350
Db 899 AAGACATGACGATACCTACCTGCTCACTGGCGTGGCAGACATCTCTTTCTCTCTA 958
QY 351 ACTCTGCCCTTTTGGGCTTATGCAAGCACTATGAGTGGGTCTTTGGCAGATCATGTGC 410
Db 959 ATTCTTCCCTTCTGGGCTACAGGAAGCCAAAGTCTCTGGATCTTTGGCGTCTACCTGTGT 1018
QY 411 AAAAATCTTCGAGGCAATGATACAACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 470
Db 1019 AAGGGCATCTTTGGCATCTATAAGTCTTTTTCAGGGGATGCTGTGCTGCTGCTGCTGCT 1078
QY 471 ATCAGAGTGTGATGCTTTTCATTGTAGTGGTCCAGGCTACCAAGGCTTCAACCGGCGAGGCT 530
Db 1079 ATCAGCATTGACCGCTAGCTAGCCATCGTCCAGCGCTGTGCGGTGTCATGCCACCGCGCC 1138
QY 531 AAGTGAAGATCTGGGCGCAAGTCATTGCTTCTCATTTGGTGGTCTCTCCCTGTTGGTT 590
Db 1139 CGCGTGTCTTCTCATCAGCAAGCTGTCTGTGGGATCTGATGCTGCTGCTGCTGCTGCTGCT 1198
QY 591 TCTTTGCCACAGATCATCTATGCGCATGTTCAAGATA-----TTGACA 633
Db 1199 TCCATCCCGAGCTGCTCTACAGCGGCTCCAGAAAGACAGCGGCGAGACAGCTGAGA 1258
QY 634 AGCTTATCTGTCAGTACCAAGTGAAGAGATATCCACTATGTTCTTTGTTATACAGATGA 693
Db 1259 TGCTCACTGGTCACTGCCCAAGT--GGAGGCTTGTATCACCATCCAAAGTGGCCAGATGG 1316
```

```
QY 694 CTCTGGGTTCTTCTGCTGCAATTTGCTACTATGATTTCTGCTACTCTCAGGCATTATCAAGA 753
Db 1317 TTTTGGGTTCCPAGTGCTATGCTGCTATGAGTTTCTGCTACCTCATTCATCCGTA 1376
QY 754 CTTGCTTCATGCTCGAAATTCAGAAAGCAGCAAAATCTTAAAGATCACTTCTCTGTAG 813
Db 1377 CTTTGTCTCCAGGCAAGCACTTTGAGCGGAACAGGCCATCAAGGTGATCATTCGCGTGG 1436
QY 814 TGGCTGTGTTCTGCTGACCCAGACACCTTTCAACC-----TTG 852
Db 1437 TGGTAGTCTTCATAGTCTTCAGCTGCGCTACATATGGGTGGTCTGCTGCTCAGACGGTGG 1496
QY 953 CCATGTTTAATCCAAAGTACAGCTGGGAGTATATACCAATCAACAGCTTTAAATATGCCA 912
Db 1497 CCAACTTCAACATCAACCAATAGCAGCTGCTGCGAAACAGCAAGCAGCTCAACATTCGCT 1556
QY 913 TCGTAGTGACAGAGGCTATAGCATCTTTGCGGCTTTCGCTTAAACCTGTACTTTATGCGCT 972
Db 1557 ATGACGTCACTCAGCGCTGCTCCGCTCGCTGCTGCTGCTCAACCTTTCTTGTATGCTT 1616
QY 973 TTGTTGGCTTAAAGTTCCGGAAGAACGCTCTGAAACTTTATGAAGGATATCGGCTGCTCT 1032
Db 1617 TCATCGGCTCAAGTTCCGAGGACCTCTTCAAGCTCTTCAAGGACTTGGGCTGCTCA 1676
QY 1033 CTCA 1036
Db 1677 GCCA 1680

RESULT 12
US-09-088-337B-23
; Sequence 23, Application US/09088337B
; Patent No. 6348574
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; Gray, Patrick W.
; Schweikart, Vicki L.
; TITLE OF INVENTION: No. 6348574e1 Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/088,337B
; FILING DATE: 01-Jun-1998
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/153,848
; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: US 07/977,452
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6348574and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEFAX: 25-3856
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2751 base pairs
```

```
;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1..691
; FEATURE:
; NAME/KEY: exon
; LOCATION: 692..1771
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 692..1768
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 2341..2348
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
; US-09-088-337B-23

Query Match 14.5%; Score 181.2; DB 4; Length 2751;
Best Local Similarity 53.2%; Pred. No. 2.3e-48;
Matches 481; Conservative 0; Mismatches 383; Indels 40; Gaps 3;

QY 171 AACAAACGCTTCCCTAAAGTTCAAGGAGGCTCTTTTGGCCCTGTGTGTACCTGGTAGTGT 230
Db 779 AGAAGGATGCGGAACCTTTAAGGCTGTGTTCTGCTCTCATGTATTCTGTCTATCTGC 838
QY 231 GTCITTTGACTGCTAGGAACTCCCTGTTCTGATTATATACATTTTCTACCAAGAGCTG 290
Db 839 TTCGTGGGCTGCTCGGCAACGGGCTGGTGATACGTACATCTATTTCAAGAGGCTC 898
QY 291 AGGACTCTGACAGATGTCTTCTGCTGAACCTTGCCCTTGGCTGACCTGGTGTGTGTCTGT 350
Db 899 AAGACCATGACGGATACCTACTCTCACTGCTGCGCTGGCAGACATCCTTTTCTCCTA 958
QY 351 ACTCTGCCCTTTTGGGCTATGCAAGCACCTATGAGTGGGTCTTTGGACAGTCACTGTC 410
Db 959 ATTCTTCCCTTCTGGGCTACAGCGAAGCAAGCTCTGGATCTTTGGGCTCTACCTGTGT 1018
QY 411 AAAACTCTTCAGGCGATGTATACAATGAATCTTACGTGTCCATGCTCACTCTCACCTGC 470
Db 1019 AAGGCACTTTTGGCATCTATAGTTAAGTCTTTCAGCGGATGCTGCTCTCTATGC 1078
QY 471 ATCAGAGTGGATCTTTTCATTTGATGGTCCAGCTACCAAGGCTTCAACCGCAGGCT 530
Db 1079 ATCAGCATTTGACCGCTACGTAGCCATGCTCCAGGCGGCTGTCGCTCATCGCCGCGCC 1138
QY 531 AAGTGAAGATCTGGGCGCAAGTCACTTTGCTGCTCATTGGGCTGCTCTCTCTTGGTT 590
Db 1139 CGCGTGTCTTCATCAGCAAGCTGTCTGTGGGCTGCTGATGCTGGCCCTCTTCTC 1198
QY 591 TCTTTGCCACAGATCATCTATGGCATTTTCAAGATA-----TTGACA 633
Db 1199 TCCATCCCGGAGCTCTCTACAGCGGCTCCAGAAAGAACAGCGCGAGGACACGCTGAGA 1258
QY 634 AGCTATCTCTCAGTACCAAGTGAAGAGATATCCACTATGGTTCTTTGTATACAGATGA 693
Db 1259 TGCTCACTGGTCAAGTGGCCCAAGT--GGAGGCGCTTGATCACCATCAAGTGCCCGAGATGG 1316
QY 694 CTCTGGGTTCTTCTGCTGCTGCTCATGATTTCTGTCTGCTACTCAGGCTATTCAGA 753
Db 1317 TTTTGGGTTCTTAGTGCTATGCTGGCTATGAGTTTCTGCTACCTCATTCATTCGTA 1376
QY 754 CTTGCTTCATGCTCGAAATTCAGAAAGCAGCAAAATCTTAAAGATCATCTTCTCTGTAG 813
Db 1377 CTTTGTCTCCAGGCAAGCACTTTGAGCGGAACAGGCCATCAAGGTGATCATTCGCGTGG 1436
QY 814 TGGCTGTGTTCTGCTGACCCAGACACCTTTCAACC-----TTG 852
Db 1437 TGGTAGTCTTCATAGTCTTCCAGCTGCGCTACATATGGGTGGTCTGCTCAGACGGTGG 1496
QY 853 CCATGTTTAATCCAAAGTACAGCTGGGAGTACTATACCAACAGCTTTAAGTATGCCA 912
```

D <sub>b</sub>	1497	CCAACTTCACATCACCAATAGCAGCTGCTGGGAAACGACGACGCTCAACATTCGCT	1556
Q <sub>y</sub>	913	TCGTAGTGACAGAGGCTATAGCATCTTTTCGGGCTTCGCTTAACCCCTGTACTTTTATGCTT	972
D <sub>b</sub>	1557	ATGAGTCACTACAGCGTGGCTCCGTCGCGTCTGCGTCAACCCCTTTCTTGTATGCTT	1616
Q <sub>y</sub>	973	TTGTTGCTTTAAAGTTCCGGAAGAACGCTCTGGAACATTATCAAGGATATCGGCTGCCTCT	1032
D <sub>b</sub>	1617	TCATCGGCGTCAAGTTCCGAGGACCTCTTCAAGCTCTTCAAGGACTTTGGGCTGCCTCA	1676
Q <sub>y</sub>	1033	CTCA	1036
D <sub>b</sub>	1677	GCCA	1680

## RESULT 13

```

PCT-US93111153-23
; Sequence 23 Application PC/RUS93111153
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweikart, Vicki L.
; TITLE OF INVENTION: Novel Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 64

```

NUMBER OF SEQUENCES: 64  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray  
ADDRESS: Bicknell  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/11153

RESIT.T 14

RESULT 14  
US-09-170-496D-175  
; Sequence 175, Application US/09170496D  
; Patent No. 6555339

GENERAL INFORMATION:  
APPLICANT: Behan, Dominic P.  
APPLICANT: Chalmers, Derek T.  
APPLICANT: Liaw, Chen W.  
TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-  
TITLE OF INVENTION: Receptors  
FILE REFERENCE: AREN-0040  
CURRENT APPLICATION NUMBER: US/09/170,496D  
CURRENT FILING DATE: 1998-10-13  
NUMBER OF SEQ ID NOS: 294  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 175  
LENGTH: 1074  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-170-496D-175

Query Match 14.3%; Score 178.6; DB 4; Length 1074;  
Best Local Similarity 52.5%; Pred. No. 8.9e-49;  
Matches 487; Conservative 0; Mismatches 404; Indels 36; Gaps 3;

QY 146 CAATCCAGTGATACAGCAGGAGACAAACGCTTCCTAAAGCTTCAAGAGGTCTTTT 205  
DB |||||  
QY 57 CAATTCAGTCTTCTACTGTGAGAAAAACAATGTCAGCAGTTTGGAGCATTCT 116  
DB |||||  
QY 206 GCCCTGTGTACTGCTAGTGTGTTGTTGACTGCTAGGAAACTCCCTGGTCTGAT 265  
DB |||||  
QY 117 CCCACCTTGTACTGGCTGTGTTTATCATCGTGGTGCTTGGCAACAGTCTTGTATCCT 176  
DB |||||  
QY 266 TATATACATTTTACAGAACTGAGGACTCTGACAGATGTTTCTGCTGACTGCTG 325  
DB |||||  
QY 177 TGCTACTGTACTGCAAGAGTGAAGACCATGACCGACATGTTCTTTGAATTTGC 236  
DB |||||  
QY 326 CCGCTGACTGTGTTGTTGCTACTCTGCTGCTTGTGGCCCTATGCGAGGACCTATGA 385  
DB |||||  
QY 237 AATGCTGACTCTCTCTTCTGCTACTCTTCCCTTCTGGCCATTGCTGCTGACCA 296  
DB |||||  
QY 386 GTGGGCTTTGGCAGCTGATGTCGAAACTCTTCGAGGATGATATACAAAGTAACTCTA 445  
DB |||||  
QY 297 GTGGAAGTTCAGACCTTCATGTGCAAGTGTGTCACAGCATGTACAAGATGAATCTTA 356  
DB |||||  
QY 446 CGTGCTCCTCACTCTCACTGCTACCTGCATCAGTGGATGTTTCATGTAGTGTCCAGGC 505  
DB |||||  
QY 357 CAGCTGTGTGTGATGATGTCATGTCATGTCGTCGAGAGTATATGCCATTGCCAGGC 416  
DB |||||  
QY 506 TACCAAGGCTTCAACCGGAGGCTAGTGGAGATCTGGGGCAAGTCAATTGTGCT 565  
DB |||||  
QY 417 CATGAGAGCACATCTTGGAGGGAGAAAGGCTTTGTACAGCAAAATGTTTGTCTTAC 476  
DB |||||  
QY 566 CATTTGGTGGTCTCCCTGTGTTCTTTGTCACAGATCATCTATGGCATGTTCAAGA 625  
DB |||||  
QY 477 CATCTGGTATTTGGCAGCTGCTCTGTCATCTCCAGAAATCTTATACGCCAAATCAAGA 536  
DB |||||  
QY 626 T-----ATTGACAAGCTTATCTGTCAGTACCAGTGGAGATATCCACTATG-- 674  
DB |||||  
QY 537 GGAATCGGCATTGCTATCTGCACCTGTTTACCCTAGCGATGAGAGCACCACAACTGAA 596  
DB |||||  
QY 675 -----GTTCTGTTTATACAGTACTGCGGGTCTTCTCCCATTTGCTACTATGAT 727  
DB |||||  
QY 597 GTCAGCTGTCTTGACCTGAGGTCATCTTCTGGGGTCTTCTTCCCTTCTGTCATGCG 656  
DB |||||  
QY 728 TCTGTGCTACTCAGGCATATCAAGACCTTGTCTGTCGTCGAAACTTCCAGAGCACAA 787  
DB |||||  
QY 657 TTGCTGTATACCATCATCATCATCACCTGATACAGCAAGAGTCTTCCAGACAA 716  
DB |||||  
QY 788 ATCTCTAAAGATCATCTCTTGTAGTGGTGTGTTCTGCTGACCCAGACACCCCTCAA 847  
DB |||||  
QY 717 AGCCAAGAAAGTGACCATCATCTCTCTGACCGCTTGTCTGCTCAGTTTCCCTACAA 776  
DB |||||  
QY 848 CCTTGGCATGTATATCAAAAGTCAAGCTGGAGTACTATACCATACACAGCTTTAAGTA 907  
DB |||||  
QY 777 CTGCATTTTGTGTGAGACCAATTGACGCTGATGCCATGTTCAATCTCCAACTGTGCCGT 836  
DB |||||

QY 908 TGCCA-----TCGTAGTGACAGAGCTATAGCATACTTTGGGCTTG 949  
DB |||||  
QY 837 TTCCACCAACATTGACATCTCTTCCAGTCCACCCAGACCATCGCTTCTTCCACAGTTG 896  
DB |||||  
QY 950 CTTAAACCTCTACTTATGCTTGTGCTTAAAGTTCGGAAGAACGCTCTGGAACCT 1009  
DB |||||  
QY 897 CTTGAACCTCTCTCTATGTTTGTGGTGAAGATCTCCGCGGATCTCTGTGAAC 956  
DB |||||  
QY 1010 TATGAAGATATCGCTGCTCTCTCA 1036  
DB |||||  
QY 957 CTTGAAGAACTTGGTTGCATCAGCCA 983  
DB |||||

RESULT 15  
US-09-299-843A-65  
Sequence 65, Application US/09299843A  
Patent No. 6107475  
GENERAL INFORMATION:  
APPLICANT: Godiska, Ronald  
APPLICANT: Gray, Patrick W.  
APPLICANT: Schweikart, Vicki L.  
TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/299,843A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/088,337  
FILING DATE: 01-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/153,848  
FILING DATE: 17-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/977,452  
FILING DATE: 17-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Jill E. Uhl  
REGISTRATION NUMBER: 43,213  
REFERENCE/DOCKET NUMBER: 27866/32059B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX:  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2085 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 177..1310  
US-09-299-843A-65

Query Match 14.2%; Score 177.8; DB 3; Length 2085;  
Best Local Similarity 52.9%; Pred. No. 2.6e-48;  
Matches 477; Conservative 0; Mismatches 387; Indels 37; Gaps 3;



This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2004, 00:36:55 ; Search time 637 Seconds  
(without alignments)  
9947.020 Million cell updates/sec

Title: US-10-603-141-1

Perfect score: 1251  
Sequence: 1 gaattcgacagaggcaga.....gtatcatggagagcgtgac 1251

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US08A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US08B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09C\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	707	56.5	2008	15	US-10-198-846-13566
2	705.4	56.4	1953	13	US-10-342-887-1299
3	705.4	56.4	1953	13	US-10-172-118-1299
4	705.4	56.4	1953	13	US-10-225-567A-347
5	704.8	56.3	1918	13	US-10-458-110-1
6	689.4	55.1	1029	9	US-09-852-156-5
7	685	54.8	1037	9	US-09-852-156-3
8	680.8	54.4	1029	15	US-10-083-168-3
9	679.2	54.3	1029	9	US-09-940-063-1
10	679.2	54.3	1029	15	US-10-174-293-1
11	679.2	54.3	1029	16	US-10-353-690-5
12	676	54.0	1029	9	US-10-083-168-64
13	648	51.8	1026	9	US-09-852-156-1
14	183.4	14.7	1074	15	US-10-251-385-23

15	183.4	14.7	2462	15	US-10-225-567A-240	Sequence 240, Appl
16	183.4	14.7	2577	9	US-09-903-377-1	Sequence 1, Appli
17	183.4	14.7	2577	9	US-09-952-385-1	Sequence 1, Appli
18	183.4	14.7	2577	10	US-09-966-755-1	Sequence 1, Appli
19	183.4	14.7	2577	14	US-10-000-759A-1	Sequence 1, Appli
20	183.4	14.7	2577	16	US-10-305-720-1095	Sequence 1095, Ap
21	180.2	14.4	1110	11	US-09-826-509-480	Sequence 480, App
22	178.6	14.3	1074	15	US-10-251-385-175	Sequence 175, App
23	177.8	14.2	1137	15	US-10-004-113-21	Sequence 21, Appl
24	177.8	14.2	2072	13	US-09-815-937-19	Sequence 19, Appl
25	177.8	14.2	2072	15	US-10-004-113-20	Sequence 20, Appl
26	169.6	13.6	30828	15	US-10-004-113-19	Sequence 19, Appl
27	163.6	13.1	1221	15	US-10-029-386-22664	Sequence 22664, A
28	163.6	13.1	2981	13	US-10-087-192-1319	Sequence 1919, Ap
29	163.6	13.1	36176	13	US-10-087-192-1318	Sequence 1918, Ap
30	162	12.9	1221	15	US-10-325-430-17	Sequence 17, Appl
31	162	12.9	1255	17	US-10-641-643-993	Sequence 993, App
32	162	12.9	1679	16	US-10-305-720-1097	Sequence 1097, Ap
33	162	12.9	2100	15	US-10-313-542-289	Sequence 289, App
34	162	12.9	3693	15	US-10-325-430-16	Sequence 16, Appl
35	161.6	12.9	1280	13	US-10-087-192-1916	Sequence 1916, Ap
36	161.6	12.9	1302	13	US-10-244-089-1	Sequence 1, Appli
37	161.6	12.9	1302	15	US-10-254-089-1	Sequence 1, Appli
38	161.6	12.9	34455	13	US-10-087-192-1915	Sequence 1915, Ap
39	161.4	12.9	1137	15	US-10-251-385-73	Sequence 73, Appl
40	161.4	12.9	1137	15	US-10-251-385-203	Sequence 203, App
41	161.4	12.9	1137	15	US-10-004-113-24	Sequence 24, Appl
42	161.4	12.9	1500	17	US-10-641-643-1465	Sequence 1465, Ap
43	161.4	12.9	2139	13	US-10-342-887-678	Sequence 678, App
44	161.4	12.9	2139	13	US-10-172-118-678	Sequence 678, App
45	161.4	12.9	2139	15	US-10-225-567A-67	Sequence 67, Appl

ALIGNMENTS

RESULT 1

US-10-198-846-13566  
; Sequence 13566, Application US/10198846  
; Publication No. US2003009974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Steimann, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-049  
; CURRENT APPLICATION NUMBER: US/10/198,846  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13566  
; LENGTH: 2008  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 2008  
; OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-13566

Query Match 56.5%; Score 707; DB 15; Length 2008;

Best Local Similarity 78.8%; Pred. No. 3.3e-225; Indels 6; Gaps 2;  
Matches 871; Conservative 0; Mismatches 231;

QY 141 TTCAACAATTCAGTGTATACAGCC---AGGAGACAAACGGTTCTTAAGTTCAAGGAG 197  
Db 144 TTTCAGAGTTTCAATTCACAGCAGCAGGAGGAGCATCAAGACTTCTTCAGTTCAGCAAG 203













; SEQ ID NO 1  
; LENGTH: 1029  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-940-063-1

```
Query Match      54.3%; Score 679.2; DB 9; Length 1029;  
Best Local Similarity 81.2%; Pred. No. 4.4e-216;  
Matches 786; Conservative 0; Mismatches 182; Indels 0; Gaps 0;  
  
QY 166 AGGAGAACAAACGGCTTCTTAAAGTTCAAGGAGGTCCTTTTGGCCCTGTGTACCTGTAG 225  
DB 62 AGGAGCATCAAGACTTCTCGAGTTCAGCAAGGTCCTTTCTGCGCTGATGTACCTGTGG 121  
  
QY 226 TGTGTGTCTTGGACTCTAGGAACCTCCCTGTTCTGATTATATACATTTCTACAGA 285  
DB 122 TGTGTGTCTGTTGCTCTGGTGGGAACCTCTGTTGCTGATCATCATCTTACCATTA 181  
  
QY 286 AGCTGAGGACTCTGACAGATGTGTTCTGCTGAACCTTGGCCCTGGGTGACCTGTGTG 345  
DB 182 AGTTGACAGGCTGACGGATGTGTTCTGTTGGAACCTTACCCTGGCTGACCTGTGTG 241  
  
QY 346 TCTGTACTCTGCCCTTTTGGSCCTATCGACGACCTTATGAGTGGTCTTTGGCACAGTCA 405  
DB 242 TCTGCATCTGCCCTTCTGGCCCTATCGACGACCTTATGAGTGGTCTTTGGCCAGTCA 301  
  
QY 406 TGTCAAAACCTCTTCGAGGATGTATACAATGAACTTCTACGTGTCCATCTCACTCTCA 465  
DB 302 TGTGCAAGGCTACTGGGCATCTACACTATTAACCTTCTACAGTCCATGCTCATCTCA 361  
  
QY 466 CCTGCATCAGAGTGGATCGTTTCAATGAGTGGTCCAGGCTACAAAGGCTTCAACCGGC 525  
DB 362 CCTGCATCAGTGGATCGTTTCAATGAGTGGTCCAGGCTACAAAGGCTTCAACCGGC 421  
  
QY 526 AGGTAAGTGGAGATCTGGGCGAAGTCACTTGGTGTCTCATTTGGTGGTCTCCCTGT 585  
DB 422 AAGCCAAAGAGATGACCTTGGGCGAAGTCACTTGGTGTCTCATCTGGGTGATATCCCTGC 481  
  
QY 586 TGTGTCTTGGCCAGATGTATCAACTGAACTTCTAGTGTCCATGCTCACTCTCACTCTCA 645  
DB 482 TGTGTCTTGGCCAGATGTATCAACTGAACTTCTAGTGTCCATGCTCACTCTCACTCTCA 541  
  
QY 646 ATACCATAACAGCTTTAAGTATGCCATCGTAGTACAGAGGCTATAGCATATCTTCGGG 705  
DB 542 GTTACCATGACGAGGCAATTTCCACTGTGTTCTTGGCCACCCAGATGACACTGGGTTCT 601  
  
QY 706 TCTGTCAATGCTCACATGATCTGTGCTACTCAGGCATTTACAGACCTTCTGCTCATG 765  
DB 602 TCTTGCACACTGCTCACCATGATGTCTGCTATTTCAGTCAATAACAAACACTGCTTCATG 661  
  
QY 766 CTCGAAACTTCCAGAACCAATCTCTAAGATCATCTTCTTGTAGTGGCTGTGTTCC 825  
DB 662 CTGAGGCTTCCAGAACCAATCTCTAAGATCATCTTCTTGTAGTGGCTGTGTTCC 721  
  
QY 826 TGTGTACCCAGACACCTTCAACCTTGGCATGTAAATCCAAAGTACAAAGCTGGAGTACT 885  
DB 722 TGTGTACCCAGATGCCCTTCAACCTTGAAGTTCATCCGACACACACTGGGAATCT 781  
  
QY 886 ATACCATAACAGCTTTAAGTATGCCATCGTAGTACAGAGGCTATAGCATATCTTCGGG 945  
DB 782 ATGCCATGACAGCTTTCACTACACCATCATGTGTGACAGAGGCTATCGCATCTCAGGG 841  
  
QY 946 CTTCGCTTACCTGTACTTATGCTTGTGCTTAAAGTTCGAGAGACCTGTGA 1005  
DB 842 CTTCGCTTAACTTGTCTCTATGCTTGTGCTGAGCTTGAAGTTCGAAAGAACTTCTGGA 901  
  
QY 1006 AACTTATGAAGGATATCGGTGCTCTCTCACCTGGAGTCTCAAGTCAATGGAAGTCTT 1065  
DB 902 AACTTGTGAAGGATGTTGTGCTCTCTTACCTTGGGCTCTCACATCAATGGAATCTT 961  
  
QY 1066 CTGAGGACAGTTCGAGACTTGTCTGCTCCCAACATGTAGACACACAGTATGTTCC 1125  
DB 962 CTGAGGACAAATTCGAGACTTTTCTGCTCCCAACATGTGAGGAGCCACAGCATGTTCC 1021
```

```
QY 1126 AATTGTAG 1133  
DB 1022 AGTTATAG 1029  
  
RESULT 10  
US-10-174-293-1  
; Sequence 1, Application US/10174293  
; Publication No. US20030165995A1  
; GENERAL INFORMATION:  
; APPLICANT: Briskin, Michael J.  
; APPLICANT: Murphy, Kristine E.  
; APPLICANT: Wilbanks, Alyson M.  
; APPLICANT: Wu, Lijun  
; TITLE OF INVENTION: No. US20030165995A1 Antibodies and Ligands for "Bonzo"  
; FILE REFERENCE: 1855.1070-001  
; CURRENT APPLICATION NUMBER: US/10/174,293  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US/09/722,064  
; PRIOR FILING DATE: 2000-11-22  
; PRIOR APPLICATION NUMBER: US 09/449,437  
; PRIOR FILING DATE: 1999-11-24  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1029  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-174-293-1  
  
Query Match      54.3%; Score 679.2; DB 15; Length 1029;  
Best Local Similarity 81.2%; Pred. No. 4.4e-216;  
Matches 786; Conservative 0; Mismatches 182; Indels 0; Gaps 0;  
  
QY 166 AGGAGAACAAACGGCTTCTTAAAGTTCAAGGAGGTCCTTTTGGCCCTGTGTACCTGTAG 225  
DB 62 AGGAGCATCAAGACTTCTCGAGTTCAGCAAGGTCCTTTCTGCGCTGATGTACCTGTGG 121  
  
QY 226 TGTGTGTCTTGGACTCTAGGAACCTCCCTGTTCTGATTATATACATTTCTACAGA 285  
DB 122 TGTGTGTCTGTTGCTCTGGTGGGAACCTCTGTTGCTGATCATCATCTTACCATTA 181  
  
QY 286 AGCTGAGGACTCTGACAGATGTGTTCTGCTGAACTTGTGCCCCCTGGCTGACCTGTGTG 345  
DB 182 AGTTGACAGGCTGACGGATGTGTTCTGTTGGAACCTTACCCTGGCTGACCTGTGTG 241  
  
QY 346 TCTGTACTCTGCCCTTTTGGSCCTATCGACGACCTTATGAGTGGTCTTTGGCACAGTCA 405  
DB 242 TCTGCATCTGCCCTTCTGGCCCTATCGACGACCTTATGAGTGGTCTTTGGCCAGTCA 301  
  
QY 406 TGTCAAAACCTTTCGAGGATGTATACAATGAACTTCTAGTGTCCATGCTCACTCTCA 465  
DB 302 TGTGCAAGGCTACTGGGCATCTACACTATTAACCTTCTACAGTCCATGCTCATCTCA 361  
  
QY 466 CCTGCATCAGAGTGGATCGTTTCAATGAGTGGTCCAGGCTACAAAGGCTTCAACCGGC 525  
DB 362 CCTGCATCAGTGGATCGTTTCAATGAGTGGTCCAGGCTACAAAGGCTTCAACCGGC 421  
  
QY 526 AGGTAAGTGGAGATCTGGGCGAAGTCACTTGGTGTCTCATTTGGTGGTCTCCCTGT 585  
DB 422 AAGCCAAAGAGATGACCTTGGGCGAAGTCACTTGGTGTCTCATCTGGGTGATATCCCTGC 481  
  
QY 586 TGTGTCTTGGCCAGATGTATCAACTGAACTTCTAGTGTCCATGCTCACTCTCACTCTCA 645  
DB 482 TGTGTCTTGGCCAGATGTATCAACTGAACTTCTAGTGTCCATGCTCACTCTCACTCTCA 541  
  
QY 646 AGTACCAAGTGGAGATATCCCACTATGTTCTTGTATACAGATCACTCTGGGTTCT 705  
DB 542 GTTACCATGACGAGGCAATTTCCACTGTGTTCTTGGCCACCCAGATGACACTGGGTTCT 601  
  
QY 706 TCTGTCCATGCTCACTATGATTTCTGTCTACTCAGGCATTTACAGACCTTCTGCTCATG 765
```



```
Db 602 TCTTCCACTGCTCACCATGATGTTCTGCTATTCTAGTCAATATCAAAACATGCTTCAAG 661
Qy 766 CTCGAAACTTCAGAAAGCACAATCTCTAAAGATCATCTTCTTGTAGTGGCTGTGTTC 825
Db 662 CTGGAGGCTTCAGAAAGCACAATCTCTAAAGATCATCTTCTTGTAGTGGCTGTGTTC 721
Qy 826 TGCTGACCAGACACCCCTCAACCTTGCCATGTTAATCCAAAGTACAAAGCTGGAGTACT 885
Db 722 TGCTGACCAGATGCCCTTCAACCTTCATGAAGTTTCATCCGAGCAGACACTGGGAATACT 781
Qy 886 ATACCATAACCAAGCTTTAAGTAGTGCATGCTAGTGCAGAGAGCTATAGCATACTTTCGG 945
Db 782 ATGCCATGACCAAGCTTTCATACACCATCATGCTGACAGAGGCCATGCAATACCTGAGG 841
Qy 946 CTGCTTAACCTGACTTATGCTTATGCTTGTGCTTAAAGTTCCGGAAGACGTCCTGGA 1005
Db 842 CTGCTTAACCTGCTCTATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 901
Qy 1006 AACTTATGAAGGATATGCTGCTCTCTCACTGGGAGTCTCAAGTCAATGGAGTCTT 1065
Db 902 AACTTGTGAGACATGCTGCTCTCTCACTGGGAGTCTCAAGTCAATGGAGTCTT 961
Qy 1066 CTGAGACAGTTCAGAGCTTCTGCTGCTCCCAATGTAGAGACCCAGTATGTTC 1125
Db 962 CTGAGACAAATCCAAAGCTTTTCTGCTCCCAATGTAGAGACCCAGTATGTTC 1021
Qy 1126 AATTGTAG 1133
Db 1022 AGTTATAG 1029

RESULT 11
US-10-353-690-5
; Sequence 5, Application US/10353690
; Publication NO. US20030215840A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Donoghue, Mary
; APPLICANT: Stagliano, Nancy
; APPLICANT: Perodin, Jacqueline
; APPLICANT: Rodriguez-Way, Amelie
; TITLE OF INVENTION: Methods and compositions for treating
; TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
; TITLE OF INVENTION: 28002, 33215, 43726, 59292, 26156, 32427, 2402, 7747, 1720,
; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
; TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
; TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 46624, 84237, 8912, 2868,
; TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
; TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
; FILE REFERENCE: MPI02-018P/RNOMIN
; CURRENT APPLICATION NUMBER: US/10/353,690
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: 60/353,224
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/364,529
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/373,861
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/376,287
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 60/388,080
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: 60/390,971
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/394,130
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/394,797
```

```
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/404,904
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/405,450
; PRIOR FILING DATE: 2002-08-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-353-690-5

Query Match 54.3%; Score 679.2; DB 16; Length 1029;
Best Local Similarity 81.2%; Pred. No. 4.4e-216; Mismatches 182; Indels 0; Gaps 0;
Matches 786; Conservative 0;

Qy 166 AGGAGAACAAACGCTTCTAAAGTTCAAGAGAGTCTTTTGGCCCTGTGTGTACCTGGTAG 225
Db 62 AGGAGCATCAAGACTTCTGCAGTTCAGCAAGGTCTTTCTGCCCTGCATGTACTGCTGGTG 121
Qy 226 TGTGTGCTTTGGAGCTAGGAAACTCCCTGCTTCTGATTATATATACATTTTCTACAGA 285
Db 122 TGTGTGCTGCTGTGGTGGGAACCTCTCTGTGTGTGTATCATCTTCTTACATA 181
Qy 286 AGCTGAGGACTCTGACAGAGATGCTTTCTGCTGAACCTTGCCTGGCTGACCTGCTGTG 345
Db 182 AGTTGCAGAGCTTGCGGATGCTTCTGCTGGAACCTACCCCTGGCTGACCTGCTGTG 241
Qy 346 TGTGTACTTGCCTCTTTGGCCCTATGACGACCTATGAGTGGGTCTTTGGCAGAGTCA 405
Db 242 TCTGCACTTGCCTCTTCTGGCCCTATGACGAGCATCCATGAATGGGTGTTTGGCCAGTCA 301
Qy 406 TGTGCAAACTCTTCGAGGCACTGTATACATGAACCTTACGTTGCTCATCTCACTCTCA 465
Db 302 TGTGCAAGAGCTTACTGGGCACTCTACACTATTAACCTTACACGTCATGCTCATCTCA 361
Qy 466 CTGTCATCAAGTGGATGCTTTTCTTGTAGTGTGTCAGGCTACCAAGGCTTCAACGGCG 525
Db 362 CTGTCATCACTGAGTGGATGCTTTTCTTGTAGTGTGTTAAGGCCACCAAGGCTCAACCCAG 421
Qy 526 AGCTGAGTGAAGATCTGGGGCCAGTCAATTTGCTGTGCTCATTTGGTGGTCTCCCTGT 585
Db 422 AGCCAAAGAGATGACCTGGGGCAAGTCAACAGCTTGTCTCATCTGGGTGATATCCCTGC 481
Qy 586 TGTGTTCTTTGCCACAGATCATCTATGGCCATGTTCAAGATATTGACAAAGCTTATCTGTC 645
Db 482 TGTGTTCTTTGCCCAAAATTTATCTATGGCAATGCTTTTAACTCTCGACAAGCTCATATGTG 541
Qy 646 AGTACCAAGTGGAGAGATATCCACTATGCTTGTGTATACAGATGACTCTGGGTTCT 705
Db 542 GTTACCATGACGAGGCAATTTCCACTGTGTTCTTGCACCCAGATGACACTGGGTTCT 601
Qy 706 TCTGTCATTTGCTCACTATGATCTGTGCTACTCAGGCATTTCAAGACCTTGTCTCATG 765
Db 602 TCTTGCACATGCTCACCATGATTTGTCTGCTATTTCAGTCATAATCAAAACACTGCTTCATG 661
Qy 766 CTGAAACTTCCAGAACACAAATCTCTAAAGATCATCTTCTGTAGTGGCTGTGTTC 825
Db 662 CTGGAGGCTTCCAGAACACAGATCTCTAAAGATCATCTTCTGCTGGTGTGCTGTGTTC 721
Qy 826 TGCTGACCCAGACACCCCTTCAACCTTGCCATGTTAATCCAAAGTACAAAGCTGGAGTACT 885
Db 722 TGCTGACCCAGATGCCCTTCAACCTCATGAAGTTTCATCCGAGCAGACACTGGGAATACT 781
Qy 886 ATACCATAACCAAGCTTTAAGTAGTGCATGCTAGTGCAGAGGCTATAGCATACTTTCGG 945
Db 782 ATGCCATGACCAAGCTTTCATACACCATCATGCTGACAGAGGCCATCGCATACCTGAGGG 841
Qy 946 CTGCTTAACCTGACTTATGCTTATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1005
Db 842 CTGCTTAACCTGCTCTATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 901
```



Query Match	51.8%;	Score 648;	DB 9;	Length 1026;
Best Local Similarity	80.8%;	Pred. No. 1.3e-205;		
Matches 805;	Conservative	0;	Mismatches 185;	Indels 6; Gaps 4;
QY	141	TTCAACAATTCCAGTGATAACAGCC---AGGAGACAACCGCTTCTCAAGTTCAAGAG	197	
DB	34	TTCAGCAGTTTCAATGACAGCAGCCAGGAGAGCATCAAGACTTCTCGAGTTTCAGCAAG	93	
QY	198	GTCTTTTGGCCCTGTGTACTCGTGTAGTGTGTCTTTGGACTGCTAGGAACCTCCCTG	257	
DB	94	GTCTTTTGGCCCTGCATGTAAGTGGTGTGTGTCTGTGGTCTGGTGGGAACTCTCTG	153	
QY	258	GTTCGTATTATATACATTTTACAGAAAGCTGAGGACTCTGACAGATGTGTTTCTGCTG	317	
DB	154	G-TGTGGTCATATCCATCTTCTACATAAGTTGCAGAGCCTGACGGATGTGTTCTCTG-TG	211	
QY	318	AACTTGGCCCTGGCTGACCTGGTGTGTTGTCTGTACTCTGCCCCCTTTTGGGCCCTATCGAGC	377	
DB	212	AACTTACCCCTGGCTGACCTGGTGTGTTGTCTGCACTCTGCCCTTCTGGGCCCTATCGAGC	271	
QY	378	ACCTATGAGTGGGTCTTTTGGCACAGTCAATGTGCAAAACTCTTCGAGGCATGTATACAATG	437	
DB	272	ATCCATGAATGGGTGTGTTGGCCAGTCAATGTGCAAGAGCCTACG-GGCATCTACACTATT	330	
QY	438	AACTTCTAGTGTCCATGCTCACTCTCACTGTCATCACAGTGGATCGTTTCAATGTPAGTG	497	
DB	331	AACTTCTACAGCTCCATGCTCATCTCACTGTCATCACTGTGATCGTTTCAATGTTAGTG	390	
QY	498	GTTCAGGCTACCAAGGCCCTTCAA CGGCAGGCTTAAGTGGAAATCTGCGGCCCAAGTCAAT	557	
DB	391	GTTAAGGCCCAACAAGGCCCTACAACAGCAAGCCCAAGAGGATGACCTTGGGGCAAGGTCAAC	450	
QY	558	TGCTGTCTCATTTGGGTGGTGTCCCTGTGTGGTGTCTTTTGGCACAGATCACTATAGGCCAT	617	
DB	451	AGCTGTCTCATCTGGGTGATATCCCTGTGTGTGTCTTGTGCCCAATATCTATGGCAAT	510	
QY	618	GTTCAGATATTTGACAAGCTTATCTGTCAGTACACACAGTGAGGAGATATCCACTATGGTT	677	
DB	511	GTCTTTAATCTCGACAAGCTATATGTGGTTTACCATGACGAGGCAATTTCCACTGTGGTT	570	
QY	678	CTTGTATTACAGATGACTCTGGGGTCTTCTCTGCGATTGTCTCACTATGATCTGTGCTAC	737	
DB	571	CTTGCCACCAGATCACATGGGGTTCCTTTTGGCCACTGTCTCAACATGATTGTCTGCTAT	630	
QY	738	TCAGGCAATATCAAGACCTTGCTTCATGCTCATGCTCGAAACTTCCAGAAAGCAAAATCTCTAAAG	797	
DB	631	TCAGTCAATAATCAAAACACTGTGTTCAATGCTGGAGCTTCCAGAAAGCAAGATCTTAAG	690	
QY	798	ATCATCTCTCTGTAGTGCTGTGTTCTGCTGAGCCAGACAGCCCTTCAACCTTGGCCATG	857	
DB	691	ATCATCTTCTGGTGTGGCTGTGTTCTGCTGAGCCAGATGCCCTTCAACCTCATGAAG	750	
QY	858	TTAATCCAAAGTACAAGCTGGGAGTATATATACCAATAACCAAGCTTTTAAGTAGTGCCTCGTA	917	

Db 417 CATGAGACATACCTTGGAGGAGAAAAGGCTTTGTACAGCAAAATGGTTTGCCTTAC 476  
Qy 566 CATTTGGGTGGTCTCCCTGTTGGTTCTTTCCACAGATCATCTATGGCCATGTTCAAGA 625  
Db 477 CATCTGGGTATGGCAGTGTCTCTGCAATCCAGAAATCTTATACAGCCAAATCAAGGA 536  
Qy 626 T-----ATTGACAAAGCTTATCTGTCACTACACAGTGGAGAGATATCCCATATG-- 674  
Db 537 GGAATCCGGCATTTGCTATCTGCACCATGTTTACCTAGCGATGAGAGCACCAACTGAA 596  
Qy 675 -----GTTCTGTTATACAGATGACTCTGGGGTTCTTCTGCAATGTTGCTACTATGAT 727  
Db 597 GTCAGCTGTCTGACCCCTGAAAGGTCAATCTGGGGTTCTTCTTCCCTTCGTGGTCAATGGC 656  
Qy 728 TCTGTGCTACTCAGGCATTAACAGACCTTGTCTCATGCTCGAACTTCAGAAAGCACA 787  
Db 657 TTGCTGCTATACCATCATCTACACCCCTGATACAGCCAGAAAGTCTTCCAGACCAA 716  
Qy 788 ATCTCTAAGATCATCTTCTTCTGTAGTGGTGTGTCTGCTGACCCAGACACCCCTTCAA 847  
Db 717 AGCCCTAAAGTGAACCATCATCTGCTGACCCGTTCTTGTCTGTCTCAGTTTCCCTACAA 776  
Qy 848 CTTTGCCATGTTAATCCAAAGTCAAGCTGGAGTACTATACCATACCAAGCTTTAAGTA 907  
Db 777 CTGCATTTTGTGTGACACCATGACGCTATGACCTATGCCATGTTCACTCCAACTGTGCCGT 836  
Qy 908 TGCCA-----TCGTAGTACAGAGGCTATACATATCTTTCGGGCTTG 949  
Db 837 TTCCACCAACATGACATCTGCTCCAGGTCAACAGACCATGCTTCTTCCACAGCTG 896  
Qy 950 CTTAAACCTGACTTTATGCTTGTGCTTAAAGTCCGGAAGAACGCTCTGGAACCT 1009  
Db 897 CTTGAACCTGCTCTCTATGTTTGTGGGTGAGAGATTCGCGGATCTCGTGA AAC 956  
Qy 1010 TATGAAGATATCGGCTGCTCTCTCA 1036  
Db 957 CCTGAAGAACTTGGTTGTCATCAGCCA 983

RESULT 15  
US-10-225-567A-240  
; Sequence 240, Application US/10225567A  
; Publication No. US20030113798A1  
; GENERAL INFORMATION:  
; APPLICANT: Lifespan Biosciences  
; APPLICANT: Brown, Joseph P.  
; APPLICANT: Burmer, Glenn C.  
; APPLICANT: Roush, Christine L.  
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS  
; FILE REFERENCE: 1920-4-4  
; CURRENT APPLICATION NUMBER: US/10/225,567A  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/257,144  
; PRIOR FILING DATE: 2000-12-19  
; NUMBER OF SEQ ID NOS: 2292  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 240  
; LENGTH: 2462  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-225-567A-240

Query Match 14.7%; Score 183.4; DB 15; Length 2462;  
Best Local Similarity 52.9%; Pred. No. 1.3e-49;  
Matches 490; Conservative 0; Mismatches 401; Indels 36; Gaps 3;  
Qy 146 CAATTCAGTATTAACAGCCAGGAGAACAAAGCTTCTTAAAGTTCAAGAGGTCTTTT 205  
Db 169 CAATTCAGTATTAACAGCCAGGAGAACAAAGCTTCTTAAAGTTCAAGAGGTCTTTT 228  
Qy 206 GCCCTGTGTACCTGGTGTGTTTGTCTTGTGACTCTAGGAAATCCCTGGTGTCTGAT 265  
Db 229 CCCACCTTGTACTGGCTGTGTTTCATCGTGGGTGCTTGGGCAACAGTCTTGTATCT 288

Search completed: September 29, 2004, 03:12:52  
Job time : 541 secs

Qy 266 TATATACATTTTCTACAGAACTGAGGACTCTGACAGATGTGTTTCTGCTGAATTCGC 325  
Db 289 TGTCTACTGTGTTACTGACAAAGTGAAGACCAAGCCAGATGTTTCCCTTTTGAATTTGGC 348  
Qy 326 CTTGGTGAACCTTGTGTTTGTCTGTACTCTGCTTCTTGGGCTATGAGGACCACTATGA 385  
Db 349 AATTGCTGACCTCTCTTCTTGTCTCTTCTTCTTGGGCAATGCTGTGCTGACCA 408  
Qy 386 GTGGGTCTTTGACACAGTCTGCAAACTCTTCGAGCATGTATACAAGAACTTTCTA 445  
Db 409 GTGAAGTTTCCAGACTTCTATGTCAAGGTGTCAACAGCATGTACAAGATGAACCTTCTA 468  
Qy 446 CTTGTCCATGCTCACTCTCACCTGCATCACAGTGGATGTTTCTATTGTAGTGGTCCAGGC 505  
Db 469 CAGCTGTGTTGCTGATCATGTGATCAGCGTGAAGGTACATTTGCCATTTGCCAGGC 528  
Qy 506 TACCAAGGCTTCAACCGGAGGCTAAGTGAAGATCTGGGGCCAAAGTCAATTTGCTTGC 565  
Db 529 CATGAGACACATACTTGGAGGAGAAAGGCTTTGTACAGCAAAATGGTTTGTCTTAC 588  
Qy 566 CATTTGGGTGGTCTCCCTGTTGGTTCTTTGCAAGATCTATGATGATGATGATGATGATG 625  
Db 589 CATCTGGGTATTTGGCAGCTGCTCTGATCCAGAAATCTTATACAGCAAAATCAAGA 648  
Qy 626 T-----ATTGACAAAGCTTATCTGACGATCAACAGTGAAGGATATCAACTATG-- 674  
Db 649 GGAATCCGGCATTTGCTATCTGCACCATGTTTACCTAGCGATGAGAGCAACCAACTGAA 708  
Qy 675 -----GTTCTGTTATACAGTACTCTGGGTTCTTCTGCTGCAATTTGCTACTATGAT 727  
Db 709 GTCAGCTGTCTTGAACCTGAAAGTCACTCTGGGGTTCTTCTTCCCTTCTGTTGCTATGGC 768  
Qy 728 TCTGTGCTACTCAGGCATTTATCAAGACCTTGTCTCATGCTCGAACTTCCAGAAAGCAAA 787  
Db 769 TTGCTGCTATACCATCATCATCTTACACCTGATACAAGCCAAAGAGTCTTCCAAAGCAA 828  
Qy 788 ATCTCTAAGATCATCTTCTTGTAGTGGTGTGTTCTGCTGACCCAGACACCTTCAA 847  
Db 829 AGCCCTAAAGTGAACCATCACTGCTGACCGTCTTTGTCTGTCTCAGTTTCCCTACAA 888  
Qy 848 CTTTGCCATGTTAATCCAAAGTACAAGCTGGAGTACTATACCATACCAAGCTTAAAGTA 907  
Db 889 CTGCAATTTTGTGGTGCAGACCATTTGACGCTATGCCATGTTTCACTCCAATCTGCGGT 948  
Qy 908 TGCCA-----TCGTAGTACAGAGGCTATAGCATACTTTTCGGGCTTG 949  
Db 949 TTCCACCAACATTTGACATCTGCTTCCAGGTCAACCCAGACCATGCTTCTTCCACAGTTG 1008  
Qy 950 CTTTAAACCTGTACTTTATGCTTGTGGCTTAAAGTTCGGAAGAACGCTCTGGAACCT 1009  
Db 1009 CCTGAACCTGTTCTCTATGTTTTTGTGGGTGAGAGATTCGCGGGATCTCGTGA AAC 1068  
Qy 1010 TATGAAGATATCGGCTGCTCTCTCA 1036  
Db 1069 CCTGAAGAACTTGGTTGTCATCAGCCA 1095